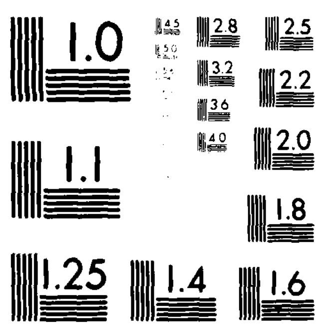
AD-A086 796 YALE UNIV NEW HAVEN CT DEPT OF PSYCHOLOGY F/G 5/10
REASONING, PROBLEM SOLVING, AND INTELLIGENCE.(U)
APR 80 R J STERNBERG N00014-78-C-0025
UNCLASSIFIED RR-3-80 NL

1 OF 2
AD
A086796

8679



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

(12)

ADA086796

Reasoning, Problem Solving, and Intelligence

Robert J. Sternberg

Department of Psychology
Yale University
New Haven, Connecticut 06520

Technical Report No. 24
April, 1980

This research was sponsored by the Personnel and
Training Research Programs, Psychological Sciences
Division, Office of Naval Research, under Contract
No. N0001478C0025, Contract Authority Identification
Number NR 150-412.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 24	2. GOVT ACCESSION NO. AD-A086 796	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 Reasoning, Problem Solving, and Intelligence		5. TYPE OF REPORT & PERIOD COVERED Periodic Technical Report, 1 Jan — 31 Mar 80
		6. PERFORMING ORG. REPORT NUMBER Research Report No. 3-80
7. AUTHOR(s) 10 Robert J. Sternberg	15	8. CONTRACT OR GRANT NUMBER(s) N0001478C0025 N00014-78-C-0025
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Yale University New Haven, Connecticut 06520		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N; RR 042-04; RR 042-04-01; NR 150-412
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, Virginia 22217	11	12. REPORT DATE 1 Apr 80
		13. NUMBER OF PAGES 173
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 14 RR-3-80 TR-24	12 192	15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

16 RR04204

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

17 RR0420401

18. SUPPLEMENTARY NOTES

To appear in R. J. Sternberg (Ed.), Handbook of Human Intelligence. New York: Cambridge University Press, in press.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reasoning, deduction, induction, problem solving, problem space, intelligence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This ~~chapter~~ report reviews the psychological literature at the interface of reasoning, problem solving, and intelligence. ~~The chapter opens~~ It begins with a discussion of some of the questions a theory of reasoning or problem solving should address, and with some suggestions as to how these questions can be addressed empirically. Next, the literature on reasoning and intelligence is reviewed. This review is divided into two parts, one on induction (over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

402628

and one on deduction; the former part pays particular attention to the literature on analogies, the latter part ~~pays particular attention~~ to the literature on linear syllogisms. Then, the literature on problem solving and intelligence is reviewed. This review, too, is divided into two parts, here, ~~one~~ on problems with well-defined solution spaces and the ~~other~~ on problems with ill-defined solution spaces; the ~~former~~ part pays particular attention to the literature on the missionaries and cannibals problem, the latter part to the literature on the hatrack problem. Finally, some general issues for the entire body of work are considered.

Reasoning, Problem Solving, and Intelligence

Robert J. Sternberg

Yale University

Chapter to appear in R. J. Sternberg (Ed.), Handbook of Human Intelligence. New York: Cambridge University Press, in press.

Running head: Reasoning & Problem Solving

Send proofs to Robert J. Sternberg
Department of Psychology
Yale University
Box 11A Yale Station
New Haven, Connecticut 06520

Reasoning, Problem Solving, and Intelligence

Reasoning, problem solving, and intelligence are so closely interrelated that it is often difficult to tell them apart. Consider, for example, the following arithmetic word problem:

> I planted a tree that was 8 inches tall. At the end of the first year it was 12 inches tall; at the end of the second year it was 18 inches tall; and at the end of the third year it was 27 inches tall. How tall was it at the end of the fourth year?

This arithmetic word problem obviously requires "problem solving" for its solution. The problem is labeled as one of "reasoning" on the test in which it appears. And the test in which it appears is one of "intelligence": the Stanford-Binet Intelligence Test (Terman & Merrill, 1937). The same fluidity of boundaries between the three constructs is equally in evidence for any of a number of problems on the Stanford-Binet. For example, "reconciliation of opposites" requires an individual to indicate in what way two opposites, such as <u>heavy</u> and <u>light</u>, are alike. In the conventional senses of the terms, reconciliation of opposites requires "reasoning," "problem solving," and "intelligence."

Whatever <u>intelligence</u> may be, <u>reasoning</u> and <u>problem solving</u> have traditionally been viewed as important subsets of it. Almost without regard to how <u>intelligence</u> has been defined, <u>reasoning</u> and <u>problem solving</u> have been part of the definition. Consider some methods for defining <u>intelligence</u>, and the roles <u>reasoning</u> and <u>problem solving</u> have played in each.

One time-honored approach to discovering the meaning of a construct is to seek expert opinion regarding the definition of that construct. The editors

of the Journal of Educational Psychology did just this in their 1921 symposium on experts' conceptions of intelligence. Almost all of the definitions provided by the 14 experts mentioned reasoning and problem solving at least implicitly. For example, Terman's (1921) definition of intelligence as "the ability to carry on abstract thinking" might be viewed as a definition of intelligence in terms of abstract reasoning, and Pintner's (1921) definition of intelligence as the "ability to adapt oneself adequately to relatively new situations in life" might be viewed as a definition of intelligence in terms of practical problem solving.

A second approach to the definitional problem might be viewed as a quantitatively more sophisticated version of the first approach. Sternberg, Ketron, Bernstein, and Conway (1979) applied factor analysis to definitional data collected from "people-in-the-street," and followed up these analyses with a comparable factor analysis of definitional data collected from experts in the field of intelligence. The motivating idea was to discover related sets of behaviors, or "factors," in people's conceptions of intelligence. Problem-solving behavior was an important factor in the conceptions of both people-in-the-street and experts.

A third approach to the problem differs from the first two in that it analyzes intelligent behavior rather than people's conceptions of intelligent behavior. The distinction between the two must be kept clear, since people's conceptions of what they do and how what they do is organized may differ from what the people actually do, and from the organization of what they actually do. This third approach, which has been widely used in the human-abilities field, is factor analysis of ability tests. Traditionally, psychometricians (specialists in psychological measurement) have sought to discover the nature of intelligence

by searching for common sources of individual-differences variation in performance on large collections of tests consensually believed to measure intelligence. Reasoning and problem solving have played important parts in virtually every theory of intelligence that has been factor-analytically derived. The earliest factor-analytic theory of intelligence, for example--Spearman's (1904, 1923, 1927)--posited a general source of individual-differences variation, _g_, common to the whole range of ability tests. Two "principles of cognition" heavily implicated in _g_, eduction of relations (e.g., "what is the relation between _lawyer_ and _client_?) and eduction of correlates (e.g., "what word-completion would result in an analogous relation from _doctor_?") are almost certainly important components of reasoning. Likewise, in Thurstone's (1938) theory of intelligence, reasoning was one of seven primary mental abilities (and in some versions of Thurstone's theory, two of eight, since inductive reasoning and deductive reasoning could be split into separate factors). Guilford's (1967) theory of the structure-of-intellect also drew heavily upon reasoning operations. Guilford's "cognition of relations," for example, appears to be essentially identical to Spearman's "eduction of relations." The importance of reasoning and problem solving to psychometric theories of intelligence is not surprising when one considers that some of the most well-known tests of intelligence comprise reasoning or problem-solving items exclusively or almost exclusively, e.g., the _Miller Analogies Test_, Raven's _Progressive Matrices_, and Cattell's _Culture Fair Test of g_.

A fourth approach, information-processing analysis, is like the psychometric approach in its application to quantitative indices of intelligent behavior (rather than to quantitative indices of conceptions of intelligent behavior), but differs from the psychometric approach in its use of stimulus variation rather

than individual-differences variation as the means to isolate elementary units of intelligence. The motivating idea in information-processing analysis is to decompose performance on tasks into elementary information-processing components, and then to show the interrelations among the components used to solve various tasks requiring intelligent performance. In this approach, too, reasoning and problem solving have been found to be critical ingredients of intelligence (Simon, 1976; Sternberg, 1977b, 1979b).

There seems to be little doubt that reasoning and problem solving play important roles in conceptions of intelligence, almost without regard to how these conceptions are derived: These roles are important as subsets of intelligent behavior. But what is the relationship between reasoning and problem solving, and even more importantly, what are they, in and of themselves? Unless we seek to stipulate the meanings of these terms from scratch, we need to look at the relationship between them in terms of the ways in which the terms have been used, and we find that the distinction between them has always been fuzzy at best. Certain kinds of problems have been studied under the rubric of"reasoning," others under the rubric of "problem solving," and it seems to be primarily an historical accident as to whether a given kind of problem has been classified as one, the other, or both. Problem solving seems to require reasoning, and reasoning seems to require problem solving. For example, it is a matter of "analogical reasoning" to complete the item, "HAPPY : ECSTATIC :: SAD : ____," but a matter of "analogical problem solving" to indicate in what ways current civilization resembles the way civilization was during the declining days of the Roman Empire.

The remainder of this chapter will be divided into four major sections. The first section will present a metatheoretical framework in terms of which theory and research on reasoning and problem solving can be understood. The next two sections will present critical reviews of the literatures on reasoning and problem solving respectively. Although the division of literature is largely arbitrary, it is nevertheless convenient. Because either of these reviews could easily require a book-length volume to do justice to the breadth of literature in each area, the emphasis in the reviews will be upon depth in the coverage of a selective subset of each literature. No attempt will be made to cover either literature in its full breadth and scope. The final section will discuss how reasoning and problem solving relate to intelligence.

A METATHEORETICAL FRAMEWORK FOR THEORY AND RESEARCH ON REASONING AND PROBLEM SOLVING

This section proposes a metatheoretical framework for theory and research on reasoning and problem solving. The section will be divided into two parts. The first will discuss the basic psychological constructs constituting the framework. The second will list and discuss questions that a theory within this framework ought to be able to answer.

Basic Psychological Constructs

The proposed metatheoretical framework is based upon the notion of the component. A component is an elementary information process that operates upon internal representations of objects or symbols (Sternberg, 1977, 1979b; see also Newell & Simon, 1972). The component may translate a sensory input into a conceptual representation, transform one conceptual representation into another, or translate a conceptual representation into a motor output. Each component has three important properties associated with it that may be measured

by mathematically (or simulatively) estimated parameters: *duration*, *difficulty*, and *probability of execution*. In other words, a given component consumes a certain amount of real time in its execution, has a certain probability of being correctly executed, and has a certain probability of being executed at all.

Components perform at least five kinds of functions (Sternberg, 1979a): *Metacomponents* are higher-order control processes that are used for planning a course of action, for making decisions regarding alternative courses of action during reasoning or problem solving, and for monitoring the success of the chosen course of action. *Performance components* are processes that are used in the execution of a reasoning or problem-solving strategy. *Acquisition components* are processes used in learning how to reason or to solve problems. *Retention components* are processes used in retrieving previously stored knowledge, whether it be knowledge needed during reasoning or problem solving, or knowledge regarding the reasoning or problem-solving algorithm itself. *Transfer components* are processes used in generalization, that is, in carrying over knowledge from one reasoning or problem-solving task to another.

Components performing each of the five kinds of functions named above can be classified in terms of three levels of generality (Sternberg, 1979a): *General components* are required for performance of all tasks within a given task universe; *class components* are required for performance of a proper subset of tasks that includes at least two tasks within the task universe; and *specific components* are required for the performance of single tasks within the task universe. A component's level of generality will depend upon the task universe under consideration: A component that is "general" in a very narrow range of tasks may be "class" in a broader range of tasks.

Questions Raised by this Framework

This metatheoretical framework suggests a number of questions that a theory of reasoning or problem-solving performance ought to be able to answer:

1. What kind or kinds of problems does the theory deal with? The answer to this question would seem to be evident from the name of a theory, but my reading of the literatures on reasoning and problem solving is that it is almost never obvious. Theorists specify only rarely either the full universe, or the subsets of the universe that are and are not covered by their theories.

How might one go about selecting tasks that are worthy of theoretical and empirical analysis? Two ways seem commonly to have been used in the past. I will summarize two ways here, show ways in which they are inadequate, and propose a third way.

First, consider the task selection procedures used by differential psychologists employing factor-analytic and other correlational techniques. Differential psychologists seem traditionally to have used either or both of two means for deciding upon what tasks to include in psychometric assessment batteries (Sternberg, 1979b).

The first means is to sample broadly from the universe of available tasks purported to measure the construct or constructs of interest. The problem with this task selection procedure is that it merely places the burden of task selection upon one's predecessors, who may have placed the burden on their predecessors, and so on. The second means of task selection used by differential psychologists is to choose tasks on the basis of their correlations with other tasks that are somehow related to the task of interest. If one selects only tasks that are perfectly intercorrelated with each other (across subjects), then the resulting tasks will probably differ from each other only trivially.

At the other extreme, choosing tasks that are uncorrelated will result in tasks having nothing in common. A more common practice in the differential literature, especially when evaluating correlations of tests with factors, has been to set an arbitrary lower limit, such as .30. In addition to such a lower limit being arbitrary, however, the limit seems to invite consideration of a plethora of tasks, some of which may be trivial variants of other tasks, and of no theoretical or practical interest in their own right. More importantly, this means of selecting tasks, and the one preceding it, lack any kind of theoretical motivation. We started off seeking a way in which theory would dictate or at least guide the selection of tasks. We have ended up with a statistical but atheoretical means of task selection that will dictate the scope of the theory. I do not wish to rule out correlational procedures entirely as an aid to task selection. But their function should be to aid rather than control task selection.

Second, consider the task-selection procedures used by information-processing psychologists using computer simulation, response time, and related procedures to understand reasoning or other psychological constructs. Newell (1973) has pointed out the dismal state of task-selection procedures among information-processing psychologists: Information-processing psychology has seemed, at times, to be more a psychology of cute tasks that have tantalized researchers than of mental phenomena in which tasks serve as a means toward understanding rather than as the end to be understood. We develop a psychology of "tasks" rather than of the mind. I do not wish to rule out tantalizing tasks from the domain of psychological research any more than I wish to rule out correlational procedures. Tasks usually maintain the interests of psychologists at least in part because they can lead to theoretically fruitful lines of research.

A task of no theoretical interest will probably last only a short time in the psychologists' toy chest. But the functional autonomy of tasks from psychological theory seems to serve no constructive purpose, and when tantalization dictates rather than aids task selection, it is serving an improper function.

Third, consider the means of task selection advocated here. In this approach, tasks are selected on the basis of four criteria originally proposed by Sternberg and Tulving (1977) in a different context: quantifiability, reliability, construct validity, and empirical validity (Sternberg, 1979a).

The first criterion, quantifiability, assures the possibility of the "assignment of numerals to objects or events according to rules" (Stevens, 1951, p.1). Quantification is rarely a problem in research on reasoning. Occasionally, psychologists are content to use subjects' introspective reports or protocols as their final dependent variable. The protocols, used in and of themselves, fail the test of quantification. If, however, aspects of the protocols are quantified (e.g., Newell & Simon, 1972) and thus rendered subject to further analysis, the quantifications of the protocols can be acceptable dependent variables so long as they meet the further criteria described below.

The second criterion, reliability, measures true-score variation relative to total-score variation. In other words, it measures the extent to which a given set of data is systematic. Reliability needs to be computed in two different ways, across item types and across subjects. Since the two indices are independent, a high value of one index provides no guarantee or even indication of a high value of the other index. Each of these two different types of reliability can be measured in two different ways, at a given time or over time.

The third criterion, construct validity, assures that the task has been chosen on the basis of some psychological theory. The theory thus dictates the choice of tasks, rather than the other way around.

The fourth criterion, empirical validity, assures that the task serves the purpose in the theory that it is supposed to serve. Thus, whereas construct validity guarantees that the selection of a task is motivated by theory, empirical validity tests the extent to which the theory is empirically supportable. The details of how empirical validation is accomplished are deferred until later.

2. What performance components are posited by the theory? A theory of reasoning or problem solving should state the performance components required for or optionally used in the solution of items of the kinds accounted for by the theory. Investigators differ, of course, in where their ideas regarding the components used come from. They may do an implicit task analysis by going through a task themselves; they may use verbal reports supplied by subjects after testing; they may use thinking-aloud protocols supplied by subjects during testing; they may use their intuitions to expand or modify previous theories.

One of the first things an investigator will want to test is whether the performance components posited by the theory to be involved in task performance are indeed used by subjects performing the reasoning or problem-solving task. A mathematical parameter can be assigned to each information-processing component in a given theory. Parameters may be of three kinds: Latency parameters represent the duration of each component; error parameters represent the difficulty of each component; probability parameters represent the probability that the component will be executed in a given task situation. Combination of each of these kinds of parameters assumes a certain kind of additivity. This assumption is testable, and to the extent it is incorrect, fits of models to data will suffer (see Sternberg, 1977b).

Response time is hypothesized to equal the sum of the amounts of time spent on each of the various components. Hence, a simple linear model predicts response

time to be the sum across the various components of the number of times each component is performed (as an independent variable) multiplied by the duration of that component (as an estimated parameter) (Sternberg, 1977a).

Proportion of response errors is hypothesized to equal the (appropriately scaled) sum of the difficulties encountered in executing each component. A simple linear model predicts proportion of errors to equal the sum across the different components of the number of times each component is performed (as an independent variable) multiplied by the difficulty of that component (as an estimated parameter). This additive combination rule is based upon the assumption that each subject has a limit on processing capacity (or space; see Osherson, 1974). Each execution of a component uses up capacity. Until the limit is exceeded, performance is flawless except for constant sources of error (such as motor confusion, carelessness, momentary distraction, etc.). Once the limit is exceeded, however, performance is at a chance level (Sternberg, 1977a).

An alternative model of item difficulty is linear with respect to logarithms of item easiness values rather than with respect to the raw easiness (or difficulty) values. In this model, the probability of answering an item correctly is equal to the product of the probabilities of performing each of the components correctly. For example, if there are two components that are theorized to be involved in performance of a task, and the probabilities of executing the two components correctly are .90 and .60 respectively, the probability of answering the problem correctly is (.90) (.60), or .54. Stated in another way, the log of the probability of answering the problem correctly is equal to the sum of the logs of the probabilities of performing each of the components correctly. Although this model of item difficulty is probably the more widely used in information-processing research, I think it is often inappropriate in the domain of

reasoning, and probably in many other domains as well. The model assumes that probabilities of erroneous (or correct) executions of components are independent across components. This assumption of independence seems rarely to be justified, however, because the probability of making an error in a component executed later during solution of a reasoning (or other) problem will generally be increased if an error was made during execution of an earlier component.

The probability of choosing a particular one of the various possible responses to a problem is assumed to be equal to the sum of the probabilities of using or combining components in each of the possible ways that can lead to that response. Obviously, the probabilities for the various responses to the problem must sum to unity.

Information-processing components can be isolated through a number of different techniques. Some of these techniques are described in Sternberg (1978b).

3. Upon what representation or representations do these components act? I doubt that there is any known test that is reasonably conclusive in distinguishing one representation for information from another. Empirical tests of alternative representations always make assumptions about information processing, since observable behavior is always the result of some set of processes acting upon some representation or representations. An information-processing model can be shown to be wrong, but can never be shown to be right, since some other information-processing model may make the same empirical predictions as a model that is not falsified. If the information-processing assumptions underlying a test of a representation are wrong, then the test of the representation is of dubious validity. But if the information-processing assumptions underlying a test are not falsified, it is still possible for the representation to be wrong, since the processing assumptions may not be correct,

or an alternative representation may exist that performs as well or better under the nonfalsified processing assumptions. At best, then, one can argue for the plausibility of a representation, but not for its ultimate correctness.

4. By what combination rule or rules are the components combined? By combination rule, I refer to the order in which components are combined, and to the use of serial versus parallel processing, exhaustive versus self-terminating processing, and independent versus nonindependent processes. The items in these latter three distinctions can be referred to as the "mode" of information processing. Order and mode apply to execution of different components, and to multiple executions of the same component.

Consider first the combination of different component processes. Suppose that two different component processes, *x* and *y*, are used in the solution of a reasoning or problem-solving item. These components may be executed in either of two (2!) different orders. Moreover, the components may be executed in various modes: First, the processes may be executed serially (*x*, then *y*) or in parallel (*x* and *y* simultaneous). If, for example, *x* is inference from *A* to *B* in an analogy of the form, *A* is to *B* as *C* is to *D*, and *y* is mapping from *A* to *C*, then either mapping may be executed immediately after inference is executed, or the two operations may be done simultaneously. When more than two processes are involved, some combination of serial and parallel processing may be used. Second, the processes may be executed exhaustively (both *x* and *y* always performed) or with self-termination (*y* executed only if execution of *x* fails to yield a solution). For example, in the analogy HE is to SHE as HIM is to (A, HERS, B. THEIRS), application of the rule that connects HE to SHE from HIM to each answer option will probably fail to yield a solution, since neither option is quite correct. But justification of one option as preferable to the

other but as nonideal will yield "HERS" as the preferred solution. In this case, justification is needed only if application fails to solve the analogy. Third, execution of each process may be independent of execution of each other process (the use or amount of use of x is uncorrelated across item types with the use or amount of use of y) or nonindependent of execution of each other process (the use or amount of use of x is correlated across item types with the use or amount of use of y). For example, if the number of attributes to be inferred from A to B of an analogy is correlated with the number of attributes to be mapped from A to C, the amounts of use of inference and mapping will be nonindependent: Larger numbers of inferences will be associated with larger numbers of mappings. Processes that are maximally nonindependent (perfectly correlated) in occurrence will be completely confounded, and hence incapable of being disentangled experimentally. Optimal distinguishability between processes occurs when their use is uncorrelated across item types. The same distinctions that apply for executions of different component processes apply as well for multiple executions of the same component process.

5. What are the durations, difficulties, and probabilities of component execution? A complete theory of human reasoning or problem solving should be able to specify not only the component processes used in reasoning or problem solving, but also the durations, difficulties, and probabilities of execution of these components. The absolute durations of various component processes are of some interest in themselves, but are of less interest than the durations of certain processes relative to certain other processes, and than the duration of a given process under a variety of experimental conditions. Durations, difficulties, and probabilities of component executions can all be estimated as parameter values via mathematical modeling or computer simulation.

6. What metacomponents are used in this form of reasoning? I have identified six metacomponents used in reasoning and problem solving (Sternberg, in press): selection of performance components for task solution, selection of one or more representations upon which these components are to act, selection of a strategy for combining the components, decision as to whether or not to maintain a given strategy, selection of a speed-accuracy tradeoff, and solution monitoring (i.e., keeping track of progress being made toward solution). Brown (1978) and Brown and DeLoache (1978) have suggested an overlapping list. Are metacomponents really needed in a theory of reasoning or problem solving? Various kinds of "meta" have become fashionable in today's research, and one might well wonder whether they are anything more than a passing fashion. Several lines of evidence suggest that "metacomponents" really are needed (see Sternberg, in press). Methods for isolating metacomponents are described in Sternberg (1979c).

7. What are the effects of (a) problem format, (b) problem content, and (c) practice upon reasoning and problem solving? Effects of problem format, content, and practice upon reasoning and problem solving can be inferred from separate internal and external validation of data for different levels of each of these variables.

Internal validation consists of the attempt to explain between-items stimulus variation in terms of an underlying model of task performance. The internal validation procedure should be applied separately to each problem format, content, and level of practice of interest. Use of these procedures for each level of each variable of interest enables one to determine specific effects of each level, for example, whether the strategy used later in practice is the same as the one used earlier during practice. External validation con-

sists of the attempt to explain between-subjects variation in terms of performance on previously validated measures that are outside the immediate paradigm of interest. The external validation procedure should be applied such that separate correlations of various scores from the experimental task and the reference tests are computed for each format, content, and level of practice. If the results of internal and external validation converge, one has a strong case for the particular argument being made. If the results diverge, alternative explanations of the obtained data must be considered.

8. What are salient sources of individual differences in reasoning or problem solving at a given age level, and how do these sources of individual differences manifest themselves? Again, there are two ways to answer this question--from the standpoint of internal validation and from the standpoint of external validation. Investigation of individual differences via internal validation is facilitated if it is possible to model individual data in just the same way that one models group data. Such modeling is usually possible for latency data if each individual contributes observations to each item data point; it is usually not possible for error and response-choice data, simply because the number of observations needed to obtain reliable probability data is prohibitive for individual subjects. If individual data are available and sufficiently reliable, one treats each individual subject as a level of a subjects variable, just as one might treat each individual item content as a level of a content variable. It is thus possible to observe what aspects of the modeling are salient sources of variation across subjects, such as the components used; the representations upon which the components act; the strategy or strategies by which components are combined; the durations, difficulties, or probabilities of execution; the consistency with which

strategies are used, etc. Investigation of individual differences via external validation involves the demonstration that identified sources of individual differences are related to patterns of individual differences in external criteria. Thus, whereas internal validation localizes the sources of variation, external validation helps interpret them and test their generalizability beyond the experimental task or tasks being investigated.

9. What are significant sources of cognitive development in reasoning or problem solving across age levels, and how do these sources manifest themselves? The sources of cognitive development, or differences across age levels, are the same as those within age levels, although the importance of various sources of individual-differences variation may be different across age levels and within age levels. In the present case, one treats the data of subjects at each age as a level of an age variable. Instructions, and sometimes the task, must be made suitable for the various age levels. For example, an analogies test that measures reasoning at a higher age level might well measure vocabulary at a lower age level.

An understanding of the development of reasoning and problem solving requires an understanding of how acquisition, retention, and transfer components operate in reasoning and problem solving tasks, and of how these kinds of components and the various other kinds of components interrelate.

In general, acquisition (retention, or transfer) of a reasoning or problem-solving skill will be facilitated by factors such as increased need for the skill, variability of reasoning or problem-solving contexts in which the skill is required, importance of the skill to solving the reasoning or problem-solving item, recency of need for the skill, helpfulness of the reasoning or problem-solving context in which the skill is required for the performance of the skill,

and helpfulness of previously stored information to the implementation of the skill in the reasoning or problem-solving situation (see Sternberg, 1979a). The importance of one or another factor will vary with the particular skill and the particular context in which the skill is required.

10. What is the relationship between a given form of reasoning and other forms of reasoning? The question posed here is one of how an investigator demonstrates communalities between tasks in the various kinds of components, and in the representations and strategies used in reasoning and problem solving. At least four tests of identity between pairs of constructs can be employed. Outcomes of these tests can suggest, but not prove, identity. First, one can demonstrate that the same information-processing model applies across tasks. Second, one can test whether values of a given parameter differ significantly across tasks. If the values do not differ, the plausibility of the argument that the parameter is the same in each task is increased. Third, one can show that any manipulation that has a certain effect upon a given component in one task has a comparable effect upon a given component in another task. Fourth, one could show that the correlation across subjects between two parameters in two tasks is close to perfect (or, in theory, to the reliability of measurement of subjects' scores).

11. What is the relationship between a given form of reasoning or problem solving and general intelligence? In order to investigate the relationship between a particular form of reasoning or problem solving and intelligence, one uses the external validation strategy described earlier for relating one form of reasoning to another. One correlates performance on the task, or the components of the task, with general intelligence as measured by some test or tests that satisfy the investigator's criteria for

an acceptable index of intelligence. (See Sternberg, 1977b, 1979a, 1979b, 1980a, for my own proposed conceptualization of intelligence.)

12. What are the practical implications of what we know about a particular kind of reasoning or problem solving? Some investigators would argue that practical implications are of no interest to them. I believe that a theory or task is of no interest if, ultimately, it bears no relation at all to practical concerns. The relation may be only tenuous at a given time, or the practical implications of a theory may be of the sort that will become clear only after a long period of time. But I do not think the issue of practical applications should be ignored altogether, lest we find ourselves studying arcane and obscure tasks that have no interest to anyone except ourselves.

Consider how the metatheoretical framework described in this section might be applied to diagnostic and prescriptive problems in educational theory and practice.

Suppose we know that a certain child is a poor reasoner. We might know this because of the child's low scores on psychometric tests of reasoning ability or because the child performs poorly in school on problems requiring various kinds of reasoning. The kinds of analyses suggested here yield a number of indices for each child (or adult) that can help localize the source of difficulty. These sources correspond to the basic sources of individual differences described above. One can discover whether certain components needed to solve one or more kinds of reasoning problems are unavailable, or available but not accessed when needed; whether the child is using a suboptimal strategy, that is, one that is time-consuming, inaccurate, or unable to yield any solution at all; whether the child finds execution of certain

components especially difficult or time-consuming; whether the child is inconsistent in his or her use of strategy; or whether the child fails in metacomponential decision-making about problem solution.

In the prescriptive domain, the first question to be addressed is whether a given information-processing strategy can be taught. One can find this out by teaching the strategy to a group of subjects, modeling the subjects' data, and determining whether the pattern of response time or error-rate data conforms to the predictions of the model. The data for each individual subject, as well as for the group, can be modeled on this basis. This kind of quantitative modeling procedure makes it possible to perform a very direct test of whether subjects have learned a particular model of information processing, in that one actually assesses exact fit between predictions and data. The fit of the trained strategy model to the data can be assessed through external validation techniques as well as through internal validation techniques. If, for example, subjects are taught to use a model of reasoning that is essentially spatial in nature, certain component scores should be theorized to correlate with scores on standard psychometric tests of spatial ability. The second question to be addressed is whether a particular model of information processing is more efficacious, on the average, than alternative models. The question can be answered simply by comparing response times and error rates under various training conditions that have been demonstrated to have been successful in imparting the proposed strategy model to the subjects. A third question is whether certain strategies are more efficacious for people with certain patterns of abilities, whereas other strategies are more efficacious for people with different patterns of abilities. This question can be answered either through cor-

relational or analysis-of-variance methodology. In the former methodology, task or component scores are correlated with scores on standard ability tests. If scores obtained using one strategy show high correlations with one kind of ability, and scores obtained using another strategy show high correlations with a different kind of ability, then one has evidence that the efficacy of a given strategy depends upon the ability pattern of the subjects using that strategy. In the latter methodology, one compares latency or error scores for subjects high and low in targeted abilities under various strategy training conditions, searching for an interaction between the strategy and aptitude patterns of the subjects. Interactions can be particularly strong when there are reasonably large proportions of subjects who are high in an ability called for by one strategy and low in an ability called for by another strategy, and vice versa.

The twelve questions posed above are obviously not the only ones that might be asked, nor are they necessarily the "right" ones that should be asked. They do seem to provide, however, a reasonable basis for testing the completeness of a theory of reasoning or problem solving falling under the general metatheoreticcal framework outlined in the first part of this section.

REASONING AND INTELLIGENCE

Reasoning may be characterized as an attempt to combine elements of old information to form new information. The old information may be external (books, magazines, newspapers, television, etc.), internal (stored in memory), or a combination of the two. The new information may be implicit but not obvious in the old information, as is the case when deductive reasoning is performed, or it may be nowhere contained in the old information, as is the

case when inductive reasoning is performed. Although it can be shown that the distinction between deduction (reasoning from given premises to a logically certain conclusion) and induction (reasoning from given premises to a reasonable but logically uncertain conclusion) is actually a fuzzy one (Skyrms, 1975), we shall maintain the distinction here as a matter of convenience, in much the same way that a distinction is maintained between reasoning and problem solving.

Inductive Reasoning

The Scope of Inductive Reasoning

In inductive reasoning, the information contained in the premises of a problem is insufficient to reach a conclusion. As a result, one can reach "inductively probable" conclusions, but not "deductively certain" ones. A number of different kinds of inductive-reasoning problems have been studied, among them:

1. Analogies, e.g., "LAWYER is to CLIENT as DOCTOR is to ?" Analogies can be composed from any of a number of different kinds of content (e.g., verbal, geometric, schematic-picture) and any of a number of different kinds of formats (e.g., fill-in-the-blank, true-false, multiple-choice). Although it is usually the last term that the subject has to induce, analogies can be presented in formats where one of the other terms is missing, or even where several of the other terms are missing (e.g., Lunzer, 1965). Reviews of the literature on analogical reasoning can be found in Dawis and Siojo (1972) and in Sternberg (1977b), as well as below. Among the original reports of theory and research on analogical reasoning are those of Ace and Dawis (1973); Achenbach (1970a, 1970b, 1971); Evans(1968); Feuerstein (1979); Gallagher and Wright (1977, 1978); Gentile, Kessler, and Gentile (1969); Gentile, Tedesco-Stratton, Davis, Lund, and Agunanne (1977); Grudin (1980); Johnson

(1962); Kling (1971); Levinson and Carpenter (1974); Lunzer (1965); Meer, Stein, and Geertsma (1955); Mulholland, Pellegrino, and Glaser (in press); Reitman (1965); Rips, Shoben, and Smith (1973); Rumelhart and Abrahamson (1973); Shalom and Schlesinger (1972); Spearman (1923); Sternberg (1977a, 1977b); Sternberg and Gardner (1979); Sternberg and Nigro (1980); Sternberg and Rifkin (1979); Tinsley and Dawis (1972); Whitely (1973, 1977, 1979a, 1979b); Whitely and Barnes (1979); Whitely and Dawis (1973, 1974); Williams (1972); Willner (1964); and Winston (1970). This set of references does not include those from the voluminous literature on matrix problems, which are similar, but not identical, to analogies.

2. Series completions, e.g., 2, 5, 8, 11, ? Series completions, like analogies, can be composed of a variety of contents (e.g., verbal, geometric, numerical, schematic-picture), and can be stated in any of a number of different forms (e.g., fill-in-the-blank, true-false, multiple-choice). Usually, the subject's task is to fill in the term following the last given one (extrapolation task), although one or more terms may be missing from the middle rather than from the end of the series (interpolation task). A review of the literature can be found in Jones (1974). Some of the original theoretical and empirical reports on series completions include those of Egan and Greeno (1974); Ernst and Newell (1969); Gregg (1967); Holzman, Glaser, and Pellegrino (1976); Jones (1971); Klahr and Wallace (1970); Kotovsky and Simon (1973); Lashley (1951); Leeuwenberg (1969); Pellegrino and Glaser (1980); Psotka (1975, 1977); Restle (1967, 1970, 1972); Restle and Brown (1970a, 1970b); Simon (1972); Simon and Kotovsky (1963); Simon and Lea (1974); Simon and Newell (1974); Simon and Sumner (1968); Sternberg (1979b); Sternberg and Gardner (1979); Thurstone (1938); Vitz and Todd (1969); and Williams (1972).

3. Classifications, e.g., "Which of the following words does not belong with the others? CAT, ELEPHANT, UNICORN, WOLF" Classifications can be presented in verbal form, or in any of the forms applicable to the other kinds of

induction problems considered above (numbers, geometric forms, schematic pictures). Although the problems are usually presented in the "odd-man-out" format used in the example, they are sometimes presented such that subjects are required to find more than one item that does not belong with the others (e.g., Cattell & Cattell, 1963), or such that subjects are required to indicate which of several answer options fits best with a set of given items (e.g., Sternberg & Gardner, 1979).

For whatever reason, the psychometric classification task has not been subject to a great deal of experimental analysis. This is surprising, since its role in the psychometric tradition has been as prominent as that of series completions, and since the processes involved in this kind of problem would seem to be of equal interest. Although the problem in its psychometric form has not been widely studied, there has been enormous interest in the psychological literature on classificatory and categorization behavior. Some perceptual approaches to this area are reviewed in Reed (1972) and some conceptual approaches are reviewed in Rosch (1977).
Some original reports that deal with the psychometric classification problem are those of Pellegrino and Glaser (1980), Sternberg (1979b), Sternberg and Gardner (1979), and Whitely (1979a, 1979b).

These three kinds of induction problems are not the only ones that have been studied, of course. A large literature exists on matrix problems (e.g., Burke, 1958; Esher, Raven, & Earl, 1942; Gabriel, 1954; Hunt, 1974; Jacobs & Vandeventer, 1971a, 1971b, 1972; Linn, 1973), as well as on causal inference (e.g., Ajzen, 1977; Carroll & Siegler, 1977; Chapman, 1967; Chapman & Chapman, 1967, 1969; Fischhoff, 1976; Gollob, Rossman, & Abelson, 1973; Kelley, 1967, 1972; Lyon & Slovic, 1976; Mill, 1843; Nisbett & Borgida, 1975;

Nisbett, Crandall, & Reed, 1976; Nisbett & Ross, 1979 ; Scriven, 1976; Smedslund, 1963; Taylor & Fiske, 1978; Tversky & Kahneman, 1974, 1977; Wason, 1968; Wason & Johnson-Laird, 1972). Other kinds of induction problems have been and might be studied as well, some of which at first glance do not even appear to be induction problems. Metaphorical comprehension, for example, can be seen as a special case of inductive reasoning (Miller, 1979; Sternberg, Tourangeau, & Nigro, 1979; Tourangeau & Sternberg, in press). For the purposes of the present review it will be sufficient to present a case study of just one kind of inductive reasoning that psychologists of all persuasions seem to agree is a critical element of intelligence, reasoning by analogy.

<u>A Case Study of Inductive Reasoning</u>: <u>The Analogy</u>

1. <u>Nature of the problem</u>. An analogy is a problem of the form <u>A</u> is to <u>B</u> as <u>C</u> is to <u>D</u> (<u>A</u> : <u>B</u> :: <u>C</u> : <u>D</u>), where, in most instances, the last term is omitted and must be filled in, selected from among answer options, or confirmed in a true-false situation. An analogy can be made arbitrarily difficult by making the terms difficult to encode. For example, the analogy, PHILOLOGY : LANGUAGES :: MYCOLOGY : (a. FLOWERING PLANTS, b. FERNS, c. WEEDS, d. FUNGI) requires only minimal reasoning ability, but is difficult because very few people know that <u>mycology</u> is the study of fungi. Analogies that derive their difficulty from the complexity of the terms rather than from the relations between terms or between relations do not necessarily measure inductive reasoning ability (see Sternberg, 1977b). Our concern here will be with analogies that derive their difficulty from their reasoning aspects rather than from their vocabulary aspects.

Performance on analogies satisfies the four criteria described in the

preceding section of the chapter. First, performance can be quantified in terms of either response latency, error rate, or distribution of responses given among the possible responses that might be given. Second, performance on analogical reasoning tasks can be measured reliably. Sternberg (1977a) reported reliabilities across items of .97 and .89 for People-Piece and geometric analogies respectively, and standard psychometric tests including sections measuring analogical reasoning typically report reliabilities across subjects in the .80's and .90's (e.g., Miller Analogies Test Manual, 1970). The construct validity of performance on tests of analogical reasoning is unimpeachable. One of the first theorists of general intelligence, Spearman (1923), used analogies as the prototypes for intelligent performance. Spearman exemplified his three basic principles of cognition through the use of the analogy. The ability to perceive second-order relations, or relations between relations, has served as the touchstone marking the transition between concrete and formal operations in Piaget's (1950) theory of intelligence, and analogies, since they require the ability to perceive relations between relations for their solution, can serve as a useful measure for distinguishing concrete-operational from formal-operational children (Sternberg & Rifkin, 1979). Finally, analogies have played a major role in information-processing theories of intelligence. Reitman (1965) and Sternberg (1977b) have used analogies as cornerstones for information-processing theories of intelligence, and other investigators have also seen analogies as fundamental to information-processing notions of intelligence (e.g., Pellegrino & Glaser, 1980; Whitely, 1977a, 1977b). Thus, analogies have played a central part in the theorizing of differential, Piagetian, and information-processing theories of intelligence. Indeed, at least two books have been written that deal almost exclusively with

analogies and their relationship to intelligence (Piaget with Montangero & Billeter, 1977; Sternberg, 1977b).

2. Performance components. All theorists seem to agree that analogical reasoners must encode analogy terms, that is, translate them into internal representations upon which further mental operations can be performed, and that these reasoners must complete analogy solution by responding with an answer to a given problem. Theorists have expressed their major disagreement over the roles of three intermediate comparison operations, called inference, mapping, and application, and over whether any additional operations need to be added to this list. Consider first the disagreements revolving around these three critical operations. We will use as an example analogy, LAWYER : CLIENT :: DOCTOR : (a. PATIENT, b. MEDICINE).

A first theory claims that inference, mapping, and application, as well as encoding and response, are all used in analogy solution. The reasoner (a) encodes the terms of the analogy, (b) infers the relation between LAWYER and CLIENT (a lawyer renders professional services to a client), (c) maps the higher-order relation between the first half of the analogy and the second (both deal with individuals who render professional services), (d) applies a relation analogous to the inferred one from DOCTOR to each answer option, choosing the correct option (a DOCTOR renders professional services to a PATIENT, not to a MEDICINE), and (e) responds (Sternberg, 1977a, 1977b). A second theory claims that only inference and application, in addition to encoding and response, are used in analogy solution. Mapping is not used (Johnson, 1962; Shalom & Schlesinger, 1972; Spearman, 1923). The various theorists use different labels for what are here called inference and application. Johnson refers to the inductive operation and the deductive operation,

Shalom and Schlesinger to the formation of the connection formula and the application of the connection formula, and Spearman to the eduction of relations and the eduction of correlates. A third theory claims that only inference and mapping, but not application, are used in analogy solution.

In this theory, mapping of the higher-order relation between the two halves of the analogy rather than application is used as the final comparison operation that determines which answer correctly solves the analogy (Evans, 1968; Winston, 1970).

Whitely and Barnes (1979) have argued that application in fact needs to be split into two subcomponents. In the first, which retains the name "application," the subject uses the relation inferred in the domain (first half) of the analogy as mapped to the range (second half) of the analogy to form a conception of the ideal solution. In the second, which Whitely and Barnes call "confirmation," the subject compares each of the answer options (in analogy formats where answer options are indeed presented) to the ideal solution. This modification was originally proposed by Sternberg (1977b, pp. 192-193) and rejected. Sternberg and Gardner (1979) have agreed with Whitely and Barnes, however, that application should be subdivided. Like Whitely and Barnes, they have referred to the construction of the ideal solution as "application;" they have referred to the comparison of each given option to the other options as <u>comparison</u>, following Sternberg (1977b).

Sternberg (1977b) has argued that an additional, optional operation needs to be added in order to complete the theories described above. This operation is one of <u>justification</u>. It is used when none of a set of presented answer options is perceived as strictly "correct." In this event, the subject justifies one answer as supperior to the other(s), although nonideal.

3. Representation of information. A wide variety of specific representations of information in analogical reasoning have been proposed by various theorists. One reason for this is that analogy-solving computer programs have been a favorite among those with interests in computer simulation and artificial intelligence (e.g., Evans, 1968; Reitman, 1965; Williams, 1972; Winston, 1970), and computer theories require a detailed specification of representation. Each computer program, of course, uses a representation that differs at least somewhat from that of other computer programs. If we consider only general classes of representations rather than specific examples of these classes, however, we find that two major classes of representations have been proposed: an attribute-value representation and a spatial representation. My current belief is that in solving analogies, subjects probably draw to some extent upon both kinds of representations and possibly other kinds of representations as well: The subjects perceive their task as one of solving analogies, and will represent information in whatever way or ways elucidate relationships between terms or between pairs of terms. If theorists of analogical reasoning can conceive of alternative ways of representing information for the solution of a given analogy, there is no reason to believe that subjects cannot do likewise.

An attribute-value representation of one kind or another has been used by all of the computer theorists, and by Sternberg (1977a, 1977b). Consider, for example, how an attribute-value representation could account for the representation of information during the solution of the analogy, WASHINGTON : 1 :: LINCOLN : (a. 10, b. 5) (Sternberg, 1977a) :

WASHINGTON might be encoded as [(president (first)), (portrait on currency (dollar)), (war hero(Revolutionary))]

1 might be encoded as [(counting number (one)), (ordinal position (first)), (amount (one unit))]

LINCOLN might be encoded as [(president (sixteenth)), (portrait on currency (five dollars)), (war hero (Civil))]

10 might be encoded as [(counting number (ten)), (ordinal position (tenth)), (amount (ten units))]

5 might be encoded as [(counting number (five)), (ordinal position (fifth)), (amount (five units))]

The attribute-value representation can be extended to pictorial as well as verbal kinds of items. A black square inside a white circle, for example, might be represented as ((shape (square)), (position (surrounded)), ((color (black))), ((shape (circle)), (position (surrounding)), ((color (white))) . The attribute-value representation can also be extended to continuous values. Terms of animal-name analogies, for example, such as TIGER in the analogy, TIGER : CAT :: WOLF : (a. ZEBRA, b. DOG), can be represented in the form, [(size (<u>x</u>)), (ferocity (<u>y</u>)), (humanness (<u>z</u>))] , where <u>x</u>, <u>y</u>, and <u>z</u> represent amounts of size, ferocity, and humanness, respectively.

A <u>spatial representation</u> of information has been used by Rumelhart and Abrahamson (1973); Rips, Shoben, and Smith (1973); and Sternberg and Gardner (1979). In each case, the domain of stimuli has consisted of animal names, although Rumelhart and Abrahamson reported that they had formulated analogies based upon terms of a color space, with equal success. The spatial representation assumes that for each term of an analogy problem, one can locate a point in a multidimensional conceptual space, and that for any analogy problem of the form A : B :: C : ?, there exists an ideal solution point in the multidimensional space that serves as the optimal completion of the analogy. It has

been found that a three-dimensional space, with dimensions of size, ferocity, and humanness, well represents a large set of mammal names (Henley, 1969).

No one has directly tested the validities of these alternative representations for information, nor is it clear how their validity could be tested directly: Representations have been assumed rather than tested.

4. Combination rules. Investigators have sought to test models predicting response latencies, response errors, and response choices. Since different combination rules have been used in each case, each will be considered separately.

Consider first the prediction of response latencies via models of real-time information processing. Sternberg (1977a) tested the three basic theories of analogical reasoning described earlier, using justification where appropriate. Application had been split into the two subcomponents in model tests for one of three experiments (see Sternberg, 1977b), but because model performance with the additional parameter clearly did not warrant addition of the extra parameter, use of comparison was discontinued. All models were assumed to be strictly linear and additive. Sternberg's (1977a) data supported the theory with all of inference, mapping, and application.

Sternberg (1977a, 1977b) compared four variants of the proposed model that differed in the order and mode of component execution. In each case, information processing was assumed to be strictly serial (mostly as a matter of convenience and simplicity), since parallel models were not tested, and executions of the various processes demanded by the problems were manipulated in order to remove significant dependencies. All variants of the basic model assumed that encoding of all attribute-values of a given term occurred in immediate succession. Models differed in which of inference, mapping, and application were exhaustive and which were self-terminating.

The data were interpreted as giving strongest support to a variant in which inference was exhaustive, and mapping and application were self-terminating. Values of R^2 were .92, .86, and .80 for People- Piece, verbal, and geometric analogies respectively. Slightly less support went to a model in which inference, mapping, and application were all self-terminating. Much less support went to two other models. A subsequent experiment with People-Piece stimuli confirmed this order of model fits (Sternberg & Rifkin, 1979).

Mulholland, Pellegrino, and Glaser (in press) also tested fits of models to latency data, in their case, for geometric analogies. Their model differed in form from those discussed above in that it separated out only encoding, transformation operations (e.g., inference), and response. A simple additive model accounted for 95% of the variance in the latency data. When an interaction term was added that multiplied the number of attribute-values to be encoded by the number of attribute-values to be compared, model fit increased slightly but significantly. The investigators argued that the interaction should be taken into account, at least for large numbers of subjects.

Sternberg and Nigro (1980) fit alternative models to latency data for verbal analogies that were constructed from a wide variety of possible conceptual relations. These investigators were particularly interested in what role, if any, word association plays in the solution of verbal analogies. With just three parameters--encoding, justification, and response (which were used for consistency in numbers of parameters across the age levels that were studied)--these investigators were able to account for 85% of the variance in their adults' latency data. Word association was found to play no significant role in the solution processes of the adult subjects.

Sternberg and Gardner (1979) fit a mathematical model to latency data obtained from subjects solving animal-name analogies. These authors were the first to use independent variables in prediction of latency data that were based upon a spatial representation (in contrast to the preceding studies, which all assumed attribute-value representations). Their model, which included only encoding, comparison, justification, and response (for comparability to other induction tasks that were studied) accounted for 77% of the variance in the latency data.

Consider next the prediction of <u>error rates</u> in analogy solution. Several investigators have sought to predict the bases for differential error rates across their various item types.

Sternberg (1977a, 1977b) used the same basic additive model to predict error rates that he had used to predict solution latencies. The only difference was in the dependent variable. Proportion of response errors was hypothesized to equal the (appropriately scaled) sum of the difficulties encountered in executing each component operation. A simple linear model predicted proportion of errors to be the sum across the different component operations of the number of times each component operation is executed (as an independent variable) multiplied by the difficulty of that component operation (as an estimated parameter). The model was successful in accounting for error rates in People-Piece and geometric analogy experiments, but not in a verbal analogies experiment. Values of R^2 were .59, .50, and .12 in the three respective experiments.

Mulholland et al. (in press) used a different model to account for their error data. These investigators claimed that their data showed independence and additivity of error probabilities associated with separately transformed

elements. Error rates can thus be understood in terms of the simple accumulation of independent, incorrect executions of information processes, any one of which leads to an error in response. The authors' logarithmic model accounted for an impressive 93% of the variance in their error data. The one thing to keep in mind in using or evaluating a model such as this one is that probabilities of errors due to different kinds of accumulated operations must be independent. Such a model would not be tenable if certain operations depended for their validity upon other earlier executed operations, e.g., application depends upon inference, and both inference and application depend upon encoding.

Consider finally the prediction of <u>response choices</u> in analogy solution. The first ones to predict response choices were Rumelhart and Abrahamson (1973). These authors used animal-name analogies, and had subjects rank-order options. They assumed that information could be represented in a multidimensional space. In order to predict response choices, Rumelhart and Abrahamson adapted Luce's (1959) choice rule to the choice situation in the analogy. The details of this choice rule need not concern us here. They further specified that the monotone decrease in the likelihood of choosing a particular answer option x_i as best follows an exponential decay function with increasing distance from the ideal point(best possible solution in a multidimensional space) of the analogy. Finally, they specified that once subjects have ranked a given alternative as first, they reapply the choice rule to the remaining alternatives to choose an option as second ranked, and continue to reapply the rules until all options have been ranked-ordered. Rumelhart and Abrahamson (1973) conducted three ingenious experiments designed to test the validity of their model of response choice in analogical reasoning. In the first experiment, they set out to demonstrate that subjects rank-order options in accordance

with the assumptions described above. In a second experiment, they tested the prediction of their model that the probability of choosing any particular response alternative X_i as the best alternative depends upon the ideal solution point and upon the alternative set, but not upon the particular terms in the analogy itself. Thus, all possible analogies with a given ideal point and alternative set should yield the same distribution of responses over the alternatives, regardless of the terms of the analogy. The data from the second experiment were somewhat consistent with the prediction. The third experiment used a concept-formation design in which subjects were required to acquire concepts for three new mammals, "bof," "dax," and "zuk." Subjects were taught these new concepts by an anticipation method. Rumelhart and Abrahamson found that after about five learning trials, subjects were able to use the imaginary mammal names in the same way that they were able to use regular mammal names in solving analogies.

Sternberg and Gardner (1979) replicated Rumelhart and Abrahamson's Experiment 1 in the context of an experiment designed to show interrelationships between various forms of inductive reasoning. Their model fits were highly comparable to those of Rumelhart and Abrahamson, providing further support for the validity of the response-choice model under the assumption of a multidimensional representation of information.

5. Durations, difficulties, and probabilities of component execution. Consider first durations of component execution. For maximum interpretability, we shall consider durations in terms of the amounts of time spent per component per problem, rather than per attribute or some other unit that depends upon the form of representation used. Consider, for example, the data of Sternberg (1977a, 1977b). Sternberg estimated latency parameters in each of his People-Piece, Verbal, and Geometric Analogy Experiments.

Estimates were for the model found to provide the best fit to the data--the one with inference, mapping,and application—and for the model variant found to provide the best fit--the one with inference exhaustive, and mapping and application self-terminating. Several aspects of the parameter estimates are worth mentioning. First, as would be expected, performance components are of longer durations for analogies with successively greater overall latencies, except for response, which is and should be approximately constant, regardless of the difficulty of the analogies. Second, encoding of analogy terms takes the greatest proportion of time in every case, even though its absolute time changes considerably with item content. Third, the amount of time spent in analyzing relations between attributes was relatively short in the People Piece Experiment (30%) and in the Verbal Analogy Experiment (29%), where simple obvious, attributes were used, but relatively long in the Geometric Analogies Experiment (57%), where the attributes were much less obvious. Finally, although the amount of time spent on response was about the same in each case, the proportion of time decreased considerably as analogy difficulty increased.

Next, consider difficulties of component execution. Sternberg (1977b) found that only self-terminating components contributed significantly to the prediction of error rates for the People-Piece analogies. In other words, errors were due for the most part to incomplete processing in self-terminating components.

Consider finally parameter estimates obtained in the prediction of the probability distribution for response choices in analogical reasoning. Rumelhart and Abrahamson (1973) on the one hand, and Sternberg and Gardner (1979), on the other, obtained similar values of α, the slope of an exponential function, for the same analogies administered to different subjects.

6. Metacomponents. Consider again the six metacomponents of reasoning and problem solving identified earlier, and what we know about each of them in the domain of analogical reasoning.

i. Selection of performance components. All of the adults I have studied in a number of experiments on analogical reasoning have been willing and able to select components of analogical reasoning from the full set described earlier. Although some subjects may select only a subset of the components in the full model for use during analogical reasoning, this seems almost certainly a matter of choice rather than of component availability: The components required for analogical reasoning seem to be readily accessible to all adults of normal mental capacity.

ii. Selection of representation(s). People seem to be able to use alternative or even multiple representations for information in analogical reasoning. Sternberg (1977b), for example, reported that an additive (overlapping) clustering representation actually provided a better fit to the group error-rate data for mammal-name analogies than did a spatial representation. In an additive clustering representation, mammals are grouped into clusters such as "rodent pests" (rat, mouse), "cat family" (cat, lion, tiger, leopard), "dog-like" (dog, wolf, fox), "wild predators" (cat, lion, tiger, leopard, wolf, fox), etc. (Shepard & Arabie, 1979). In the Sternberg (1977b) data, the additive clustering model proposed accounted for 56% of the variance in the error data, whereas the spatial model proposed accounted for only 28%. Michael Gardner and I have replicated this finding for the same analogies but for different subjects in unpublished data we have collected. Nevertheless, the spatial representation fits response-choice data extremely well, and it is not even clear how the

additive clustering representation could be applied to data of this sort. I am inclined to regard the two kinds of modeling as elucidating different aspects of subjects' reasoning about the mammal names. If the results of fitting both kinds of representations are sensible, then it is quite possible that both are "correct": They elucidate different aspects of the ways in which subjects conceive of relations between elements in a given data set.

iii. Selection of strategy for combining components. Many of the most important theoretical questions about analogical reasoning concern strategy selection, although very few of them have yet been answered. Consider, for example, the question of whether subjects can use inherent properties of analogies to simplify their processing of information when the need arises. Consider, for example, the analogy, SNOW : BLOOD :: WHITE : RED. The models described earlier would all call for inference of the relation between SNOW and BLOOD; some of these models would then call for mapping of the relation between SNOW and BLOOD on the one hand, and between WHITE and RED on the other. But one important property of analogies is that as proportions, they can be viewed as relating (A,C) to (B,D) as well as they can be viewed as relating (A,B) to (C,D). In the above analogy, certainly it would be to the subject's advantage to infer the relation between SNOW and WHITE, and then to map the relation between SNOW and WHITE on the one hand and between BLOOD and RED on the other. Sternberg (1977b, pp. 232-233) tested this model variant for verbal analogies, and found a slight improvement in fit over that obtained for the standard strategy, suggesting that subjects might flexibly alter their strategy such that if the semantic relation between the first and third terms is closer than that between the first and second, then they infer the relation between (A,C)--the first and third terms--and map the relation

between (A,C) and (B,D). Grudin (1980) has presented even stronger data arguing for this flexibility in strategy.

iv. Decision as to whether to maintain a strategy. We currently have no evidence that subjects change their strategy for analogy solution with practice. Hence, as far as we know, subjects generally do decide to maintain whatever strategy they have started with.

v. Selection of a speed-accuracy tradeoff. At the present time, we know that a micro-tradeoff between speed and accuracy is found in analogical reasoning: Greater speed is attained at some cost in accuracy (Sternberg, 1977b). Data currently being collected by Miriam Schustack and myself also indicate that is is possible to induce a macro-tradeoff between speed and accuracy, at least when the analogies are presented tachistoscopically. In other words, different instructions regarding the relative importance of speed and accuracy can result in differential speed-accuracy tradeoffs. More interesting than the question of whether a tradeoff can be induced, however, is the question of where the loci of the tradeoff reside. Schustack and I seek to discover these loci in our experiment.

vi. Solution monitoring. We don't yet know just how subjects monitor their solution of analogy problems, but we know that they do monitor their performance. This monitoring manifests itself through the execution of the justification component, which is executed when none of the response options fit the subject's ideal conception of what the answer to an analogy should be. Justification seems to take the form of checking and possibly re-executing previously executed operations. Subjects seek to find either errors of commission (an operation was executed incorrectly) or errors of omission (an operation was executed incompletely, resulting in failure to encode an important attribute or to conceive of an important relation).

7. *Problem format, problem content, and practice.* All three of these variables have been found to affect analogical reasoning performance. Consider first the effects of *problem format*.

Some investigations have been based upon true-false analogies (Ingram, Pellegrino, & Glaser, 1976; Mulholland et al., in press; Sternberg, 1977a, 1977b). The format is of questionable merit when the analogy attributes are nonobvious and ill-defined, as in the verbal analogies of Ingram et al. (1976) and of Sternberg (1977a, 1977b). I now believe the format may be inappropriate for such analogies, since it is not clear that for analogies with ill-defined attributes, any one completion is strictly correct (at least for the terms available in the English language). For example, is the analogy, WHITE : BLACK :: BIG : SHORT true or false? In some senses of the words "big" and "short," it is true; in others, it is false. Ingram et al. found that analogies with completions that are "near misses" had longer latencies than analogies that were either "true" or obviously false, and they proposed a model that could handle this finding. Other investigators have used two-option forced-choice analogies (Sternberg, 1977a, 1977b for geometric analogies; Sternberg & Rifkin, 1979). Sternberg (1977b) has described two "extreme" strategies that might be used in scanning options in the forced-choice format. Still other investigators have used four-option multiple-choice analogies (Rumelhart & Abrahamson, 1973; Sternberg & Gardner, 1979; Whitely & Barnes, 1979). The strategies subjects use in solving problems of this nature seem to be very complex.

Lunzer (1965) presented analogies that had either just one term missing (which could be either A, B, C, or D), or that had two terms missing (either C and D, A and B, B and C, or A and D). As would be expected, analogies with a

single term missing were easier to solve than analogies with two terms missing. Those with C and D missing or with A and B missing were found to be less difficult than those with B and C missing or with A and D missing. Presumably, this was because for the latter two kinds of problems, there was no single relation linking the missing pair: Each of the two missing terms was involved in a different relation within the analogy.

Levinson and Carpenter (1974) presented verbal problems in two forms--as analogies and as quasi-analogies. An analogy took the form exemplified by "Bird is to air as fish is to ____." A quasi-analogy took the form exemplified by "A bird uses air; a fish uses ______." The quasi-analogies thus supplied the relationship, whereas the analogies did not. There were no significant differences in performance on the two types of problems for the oldest subjects tested, who were 15 years old (but see section on between-age differences that follows).

Sternberg and Nigro (1980) presented analogies to subjects in each of the following three forms:

1. NARROW : WIDE :: QUESTION : (TRIAL, STATEMENT, ANSWER, ASK)

2. WIN : LOSE :: (DISLIKE : HATE), (EAR : HEAR), (ENJOY : LIKE), (ABOVE : BELOW)

3. WEAK : (SICK :: CIRCLE : SHAPE), (STRONG :: POOR : RICH), (SMALL :: GARDEN : GROW), (HEALTH :: SOLID : FIRM)

Numbers of answer options ranged from two to four. Adult subjects were fastest on the first form and slowest on the last form (but see section on between-age differences that follows). However, error rates were highest for the second form. Adult subjects solved the problems exhaustively, in the sense that they passed through all of the answer options sequentially before selecting a response.

Johnson (1962) used what he called a method of "serial analysis" in his presentation of verbal analogies. Analogies were presented tachistoscopically, with trials divided into two parts. In the first part, subjects received the first half of the analogy. They had as long as they wanted to view this half of the problem, and when they were done studying it, they initiated the second part of the trial. In this part of the trial, subjects received the second half of the analogy. Subjects terminated this part of the trial by responding to the analogy as quickly as possible. Response latency was longer for the second half of the trial (mean latency = 6.68 seconds) than for the first half of the trial (mean latency = 3.33 seconds).

Sternberg (1977b) extended Johnson's method of serial analysis in a method of "precueing." Analogies were presented in two parts. The first part could consist of either no terms, one term, two terms, or three terms. The second part of the trial always consisted of the full analogy. Sternberg found that cue times--latencies for the first part of the trial--increased with greater amounts of precueing in the first part of the trial, and that solution times--latencies for the second part of the trial--decreased.

Consider now effects of problem content. Sternberg (1977a, 1977b) found verbal analogies to be more difficult than People-Piece analogies, and geometric analogies to be more difficult than verbal analogies. This ordering probably said more about the particular instantiations of content Sternberg used than about intrinsic properties of the content, since no attempt was made to equate concepts in any way. As mentioned earlier, for example, verbal analogies can be made arbitrarily difficult by using terms that pose vocabulary or general information demands beyond the capacities of many subjects solving the problems. Tests such as the Miller Analogies Test are difficult largely because the vocabulary and general information demands are so high (Sternberg, 1977b, 1978).

Tinsley and Dawis (1972) did an experiment that specifically set out to equate conceptual difficulty of items with two different kinds of content. In their items, the same objects were presented verbally and figurally, so that the only difference between the two sets of analogies they studied was in the content vehicle through which the objects were expressed. These authors found no significant difference in difficulty between the two contents, and they also found a correlation of .86 between the 30 items constituting each form of the test. Unfortunately, all subjects received both types of analogies in immediate succession, with the verbal analogies always coming first. As a result, it is difficult to know to what extent the results were affected by the fact that each subject received each item twice in rapid succession, and by the fact that order of presentation and item content were confounded.

Johnson (1962) presented items that were intended to be difficult either because of vocabulary demand in the first half of the analogy or because of vocabulary demand in the second half of the analogy. For example, FELINE is to CANINE as CAT is to ? was presumed to be difficult because of the vocabulary demand in the first half of the item, whereas LOSE is to WIN as LIABILITY is to ? was presumed to be difficult because of the vocabulary demand in the second half of the item. In all three item formats--response production, multiple-choice, and multiple-choice with options containing only the first letter of the option--items were more difficult (higher response latency and higher error rate) if the greater vocabulary demand was in the first half of the trial. (Recall that Johnson divided his trials into two parts, a part in which the first half was presented and a part in which the second half was presented.) Johnson found that more time was spent on the first part of the

trial for items where the vocabulary load was in the first part than for items where the vocabulary load was in the second part; he found that more time was spent on the second part of the trial for items where the vocabulary load was in the second part than for items where the vocabulary load was in the first part.

Consider finally the effects of practice upon analogy solution. Sternberg (1977b) compared performance during a first session of People-Piece analogy solution to performance during a fourth (and final)session. As would be expected, latencies and error rates decreased from the first session to the fourth. All components showed shorter latencies during the fourth session than during the first except for inference. There was no evidence of strategy change across sessions: Fits of the various models and variants of models were almost identical in the two different sessions. The most interesting difference in results showed up during external validation of scores: In the first session, no correlations of latencies for the second (solution) part of the trial with reasoning tests were significant; in the fourth session, more than half of the correlations were significant, and many of them were of high magnitude, reaching into the .60s and .70s. Sternberg (1977b)noted that this pattern of difference in the correlations is related to previous findings. Noble, Noble, and Alcock (1958) used tests from the Thurstone Primary Mental Abilities battery to predict individual differences in trial-and-error learning. They found that prediction was higher for total correct scores than for initial correct scores. These data suggested that the higher correlations resulted from performance during later trials. Fleishman and Hempel (1955) and Fleishman (1965) found that the percentage of variance accounted for in motor tasks by traditional psychometric tests increased with practice on the motor tasks. These results led Glaser (1967) to conclude that psychometric test scores are more highly correlated with performance after asymptote is reached than with performance during initial trials of practice.

Vygotsky (1962) noted that mental testing is usually based upon performance on tasks for which no explicit training has been given. He suggested that it might be more appropriate to measure performance after training and practice rather than before, at the upper rather than at the lower threshold of performance, because "instruction must be oriented toward the future, not the past" (p.104). The present data are consistent with Vygotsky's notions, and with Ferguson's (1954) notion of intelligence as "performance at some crude limit of learning" (p. 110).

8. Individual differences within age level (adults). Sternberg (1977b) found substantial individual differences in the speeds at which the various performance components of analogical reasoning are executed, and in the degree to which subjects used any systematic strategy at all. No substantial individual differences were found in components or forms of representation used, and no one else has found such individual differences either. Sternberg also failed to find clearcut individual differences in strategies for combining components, other than that some people appeared to use the third variant of the model described earlier (inference exhaustive, mapping and application self-terminating) and that other people appeared to use the fourth variant (inference, mapping, and application self-terminating).

Whitely and Barnes (1979), on the other hand, have interpreted data they have collected as evidencing important differences in strategy among adults. These authors used a "simulation task" to study verbal analogical reasoning. In this task, analogies were composed of pronounceable nonwords, e.g., LYOMON : FIRMANI :: DULCIVER : (BANSHER, PONTO, NAX, SQUISH). ("Squish" is actually a word, although it was one of their stimuli.) The terms were supposed to refer to animals that could evolve on other planets. In the

simulation task, subjects could request information about any of the analogy terms in any order. Information consisted of a picture of the animal and a list of five properties of the animal. Whitely and Barnes's results suggest that substantial strategy differences may occur in analogical reasoning. Some unpublished data monitoring subjects' eye movements during analogical reasoning have been collected by Richard Snow at Stanford, and these data also suggest strategy differences. The Whitely-Barnes data, however, are still inconclusive. First, it is not yet clear whether what subjects do during the simulation task actually "simulates" what they do during normal solution of verbal analogies. Both Reitman (1965) and Sternberg (1977b) found that subjects had only the foggiest idea of how they went about solving verbal analogies, and the fact that scores on Whitely and Barnes's simulation task correlated only .20 ($p < .05$) with scores on a psychometric verbal analogies test might lead one to question whether the two tasks do indeed draw upon the same strategies and other elements of reasoning. In drastically reducing their rate of work, subjects may also be changing their style of work. Second, the artificiality of the stimuli may lead to specialized strategies not applicable to ordinary analogy problems. The way information is encoded about a "squish" may or may not correspond to the way in which it is encoded about a real animal. Whether or not the simulation task measures the same kind of reasoning as the standard analogical reasoning task, the form of reasoning it does measure seems to be potentially interesting in its own right, combining as it does the need for learning of concepts and for reasoning with those concepts.

9. Differences across age levels. A great deal of developmental work has been done in the domain of analogical reasoning, and it is possible to touch upon only a fraction of that work here.

Piaget, with Montangero and Billeter (1977), has suggested three stages in the development of reasoning by analogy. Understanding of these stages requires some knowledge of the paradigm these investigators used to study the development of analogical reasoning. The investigators presented 29 children between the ages of 5 and 13 with sets of pictures. They asked the children to arrange the pictures into pairs. The children were then asked to put together those pairs that went well together, placing groups of four pictures into 2 x 2 matrices that represented relations of analogy among the four pictures. Children who had difficulty at any step of the procedure were given prompts along the way. Children who finally succeeded were presented with a countersuggestion to their proposed solution, by which the investigators hoped to test the strength of the children's commitment to their proposed response. At all steps along the way, children were asked to explain their reasons for grouping things as they did. In the first proposed stage of Piaget's model, characterizing the performance of children of ages 5 to 6, children can arrange pictures into pairs, but the children ignore higher-order relations between pairs. Thus, although these children can link A to B or C to D, they cannot link A-B to C-D. In the second stage, characterizing the performance of children from about 8 to 11 years of age, children can form analogies, but when challenged with countersuggestions, they readily rescind their proposed analogies. Piaget interprets this finding as evidence of only a weak or tentative level of analogical reasoning ability. In the third stage, characterizing the performance of children of ages 11 and above, children form analogies, are able to state explicitly the conceptual bases of these analogies, and resist counter-suggestions from the experimenter.

Lunzer (1965) presented children of ages 9 through 17+ with verbal analogies taking the various forms described earlier (e.g., just the A term missing, just the C term missing, the A and B terms missing, the A and D terms missing). Lunzer found that children had great difficulty with even the simplest analogies until about 9 years of age, and did not show highly successful performance until the age of 11. Lunzer concluded that even the simplest analogies require recognition of higher-order relations that are not discernible to children who are not yet formal-operational. The suggestion of three stages in Lunzer's work (before age 9, between ages 9 and 11, after age 11) seems correspondent to the suggestion of Piaget regarding three stages of reasoning by analogy.

Gallagher and Wright (1977, 1978) have done research comparing the relative abilities of children in grades 4 to 6 to provide what these investigators called "symmetrical" or "asymmetrical" explanations of analogy solutions. Symmetrical explanations showed awareness of the higher-order relation linking A-B to C-D. Asymmetrical explanations ignored this relation, dealing either only with the C-D relation, or with both the A-B and C-D relations, but in isolation from each other. Percentages of symmetrical responses increased with age and were associated with higher level of performance on the analogies.

Levinson and Carpenter (1974) presented verbal analogies (e.g., "BIRD : AIR :: FISH : ?") and quasi-analogies (e.g., "A bird uses air; a fish uses ?") to children of ages 9, 12, and 15 years of age. The standard analogies required recognition of the higher-order analogical relationship; the quasi-analogies essentially supplied this relationship. The investigators found that whereas nine-year-olds could answer significantly more quasi-analogies than analogies correctly, twelve and fifteen-year-olds answered approximately

equal numbers of each kind of item. Moreover, whereas performance on the standard analogies increased monotonically across age levels, performance on the quasi-analogies did not increase. These results provide further evidence for the ability of formal-operational children, but not concrete-operational children, to use second-order relations in the solution of verbal analogies.

Sternberg and Rifkin (1979) investigated the development of analogical reasoning processes with two kinds of schematic-picture analogies. One kind was the People-Piece analogy used by Sternberg (1977a, 1977b); the other kind was also a schematic figure of a person. But whereas the People Piece analogies were composed of perceptually integral attributes, the schematic-figure analogies were composed of perceptually separable attributes (see Garner, 1974). In the perceptually integral attributes, the levels of one attribute depend upon levels of other attributes for their existence. For example, in the People Pieces, a level of height cannot be represented without representing some level of weight (two of the attributes), and vice versa. In the perceptually separable attributes, levels of attributes are not dependent in this way. For example, in the schematic figures, the color of a hat can be represented without representing the type of footwear a figure has on. For analogies with perceptually integral attributes, subjects became more nearly exhaustive in their information processing with increasing age. This tendency to become more nearly exhaustive in information processing appears to be a general higher-order strategy in cognitive development (Brown & DeLoache, 1978), and appears to be associated with dramatic decreases in error rate over age (Sternberg,

1977b; Sternberg & Rifkin, 1979). Moreover, it was found that although fourth-graders, sixth-graders, and adults solved the analogies by mapping the higher-order relation between the two halves of the analogies, second-graders did not. Once again, then, we have support for a late developing ability to recognize and utilize higher-order relations. For the analogies with perceptually separable attributes, subjects at all ages used the same fully self-terminating strategy, the same one that second-graders used for the perceptually integral attributes. Mapping was not used, or was used for a constant amount of time across item types. (The two outcomes are experimentally indistinguishable.) Thus, it appears that for analogies with integral attributes, strategy changes with age, whereas for analogies with separable attributes, it does not change.

Some investigators have been particularly interested in the role word association plays in children's solution of analogies. The pioneering studies in the role of association in analogy solution were done by Achenbach (1970a, 1970b, 1971), who found that children in the intermediate and early secondary school grades differ widely in the extent to which they use word association as a means of choosing one from among several response options. Moreover, the extent to which children use word association serves as a moderator variable in predicting classroom performance: Correlations between performance on IQ tests and school achievement were substantially lower for children who relied heavily on free association to solve analogies than for children who relied primarily on reasoning processes. Gentile, Tedesco-Stratton, Davis, Lund, and Agunanne (1977) further investigated children's associative responding, using Achenbach's CART (Children's Associative Respond-

ing Test). They found that associative priming can have a marked effect on test scores, leading children either toward or away from correct solutions. Sternberg and Nigro (1980) found that third and sixth graders used word association to a significant extent in the solution of verbal analogies, but that ninth graders and adults did not.

Sternberg and Nigro (1980) were interested not only in whether children use word association or not, but also in how they use it in analogy solution. They found that in analogies with the three formats they used (described earlier), word association is used to guide search among options, with higher association responses being examined before lower association ones. The ninth graders and adults, however, did not use word association in this way, or in any other way that was discernible from their data.

10. Relationships between analogical reasoning and other kinds of reasoning. Several investigators have asserted that the processes used in solving analogies are used as well in solving other kinds of induction problems, such as series completions and classifications (Greeno, 1978; Pellegrino & Glaser, 1979, 1980; Sternberg, 1977b, 1979c, 1979d). Sternberg and Gardner (1979) conducted two experiments designed to test this assertion. These experiments used animal-name analogies (including those of Rumelhart & Abrahamson, 1973, e.g., TIGER : CHIMPANZEE :: WOLF : (a. RACCOON, b. CAMEL, c. MONKEY, d. LEOPARD)); series completions (e.g., SQUIRREL : CHIPMUNK : (a. RACCOON, b. HORSE, c. DOG, d. CAMEL))where the subject's task was to indicate which of four completions would follow next in a series; and classifications (e.g., ZEBRA, GIRAFFE, GOAT, (a. DOG, b. COW, c. MOUSE, d. DEER)),where the subject's task was to indicate which of the options fit best with three terms in the item stem.

In a first experiment, subjects were asked to rank-order each of four answer options in terms of goodness of fit. Correlations were .99 between the data sets for analogies and series completions, .97 between the data sets for analogies and classifications, and .98 between the data sets for series completions and classifications. The fits of the exponential model to each of the data sets were also very good: $r = .97$ for analogies, $r = .98$ for series completions, and $r = .99$ for classifications. Moreover, parameter estimates were quite similar across the three tasks. The results suggest communalities in decision rules for response choice across the three inductive reasoning tasks. The second experiment of Sternberg and Gardner used the same items, except that the number of answer options was reduced from four to two. Subjects were nonoverlapping with those in the first experiment. The main dependent variable in this experiment was solution latency rather than response choice. The authors fit a four-parameter model to data sets from each of the three tasks. The model fit the data for each task reasonable well: $r = .88$ for analogies, $r = .82$ for series completions, and $r = .78$ for classifications. Only the estimated value of the justification parameter differed significantly across tasks. These data were interpreted as providing further support for a unified model of performance in the three inductive reasoning tasks.

In a new and innovative approach to modeling cognitive abilities, Whitely (1979b) fit a three-component latent trait model to data for a verbal analogy and a verbal classification test (see also Whitely, 1979a). Whitely (1979c) then used this latent trait model as a basis for covariance modeling (Jöreskog, 1969, 1970; Jöreskog & Sörbom, 1978) intended to account for individual dif-

ferences in analogy and classification performance, as well as differences in item difficulties. The same basic model provided a good fit to data for both the verbal analogies and the verbal classifications tasks, although parameter estimates differed.

Further support for the possible unity of performance components in at least some induction tasks is provided by the success of computer programs such as William's (1972) Aptitude-Test Taker and Simon and Lea's (1974) General Rule Inducer (GRI) program, both of which can solve induction problems of a variety of types. The success of these program shows that a single set of processes can be sufficient for solving various kinds of induction problems, although of course it does not show that people actually use a single set of processes.

11. Relationship between analogical reasoning and intelligence. Theorists such as Spearman (1923), Piaget (1950, Piaget, with Montangero & Billeter, 1977), and Reitman (1965) have argued for many years that intelligence and analogical reasoning are intimately connected. Spearman's three principles of cognition were based upon three processes Spearman believed to be involved in reasoning by analogy; Piaget's concrete- and formal-operational stages are differentiated in children by the children's ability to solve analogies; and Reitman's theory of intellectual functioning is based upon solution of analogies. Moreover, empirical data collected by these investigators and others have been consistent with the strong claims made regarding the centrality of analogical reasoning in intelligent functioning. Factor-analytic investigations such as those conducted by Spearman (1904, 1927), for example, have consistently shown analogy items to be among the

highest in their loadings on *g*, or general intelligence. Some recently collected data have provided further insights into the information-processing relationships between analogical reasoning and intelligence.

Sternberg (1977b) found that each of the major "reasoning operations" in analogical reasoning--inference, mapping, application, justification--can correlate with performance on tests of general intelligence when the attributes of the analogies being solved are nonobvious. As would be expected, faster component execution was associated with higher scores on the psychometric tests. Sternberg also found an association between encoding speed and measured intelligence, except that the association went the opposite way: Slower encoding was associated with higher measured intelligence. Sternberg interpreted this finding in metacomponential terms: Brighter subjects spend relatively longer in encoding the stimuli in order to facilitate their execution of subsequent performance components that will draw upon the results of the encoding. Evidence from his studies supported this interpretation. Even the response constant shows a strong correlation with measured intelligence: Faster speed of response was associated with higher measured intelligence. This basic finding has since been replicated in the analogies task (Mulholland et al.,) and in other tasks as well (Egan, 1977). This finding may have a metacomponential interpretation also: The response constant includes confounded within it all sources of response variation that are constant across items. Strategy planning, solution monitoring, and other metacomponents are likely to be constant in duration across item types, and hence to be confounded with the response constant. These sources of variation may be responsible for the correlation

of the response constant with IQ. Sternberg (1979c) has gone so far as to suggest that the metacomponents underlying the variation are the most important elements to be reckoned with in a theory of intelligence. Finally, Sternberg (1977b) found that for verbal analogies, at least, greater systematicity in use of a strategy in solving analogies was associated with higher intelligence. All of these findings have received developmental confirmation to support the data obtained with adults. All component latencies show a general decrease with age, except for encoding, which after an initial decrease (due, presumably, to cognitive development in encoding skills) shows a subsequent increase in latency (Sternberg & Rifkin, 1979); and systematicity in strategy utilization also increases with age, at least for some types of analogies (Sternberg & Nigro, 1980; Sternberg & Rifkin, 1979).

Whitely (1976, 1977) has also shown significant relationships between measured intelligence and performance in analogical reasoning. In her 1976 study, she found significant multiple correlations (in the .30s, .40s, and .50s) between scores on verbal analogies expressing different semantic relations and scores on standard measures of mental ability. In her 1977 study, she found correlations ranging up to the .60s between numbers correct on an analogies test and scores on the Differential Aptitude Test and the Lorge-Thorndike Intelligence Test. Correlations for latencies of analogy solution were somewhat lower, but still statistically significant for items that were answered correctly.

To summarize, the history of theory and research on relationships between analogical reasoning and intelligence shows the two to be strongly related. Indeed, according to some theorists, an understanding of intelli-

gence requires an understanding of analogical reasoning: Spearman (1923), Reitman (1965), and Sternberg (1977b) all based parts of their theories of intelligence on their theories of reasoning by analogy.

12. Practical relevance. Sternberg (1977b) has argued that reasoning by analogy is pervasive in everyday experience. "We reason analogically whenever we make a decision about something new in our experience by drawing a parallel to something old in our experience. When we buy a new pet hamster because we liked our old one or when we listen to a friend's advice because it was correct once before, we are reasoning analogically" (p. 99).

Oppenheimer (1956) has pointed out the signal importance of analogy in scientific reasoning of the kind done by scientists and even nonscientists on an everyday basis:

> Whether or not we talk of discovery or of invention, analogy is inevitable in human thought, because we come to new things in science with what equipment we have, which is how we have learned to think, and above all how we have learned to think about the relatedness of things. We cannot, coming into something new, deal with it except on the basis of the familiar and old-fashioned. The conservatism of scientific enquiry is not an arbitrary thing; it is the freight with which we operate; it is the only equipment we have. (pp. 129-130)

Analogical reasoning also plays an important role in legal thinking, where it may be called "reasoning by example" (Levi, 1949):

> The basic pattern of legal reasoning is reasoning by example. It is reasoning from case to case. It is a three-step process described by the doctrine of precedent in which a proposition descriptive of the first case is made into a rule of law and then applied to a next similar situation. The steps are these: similarity is seen between cases; next the rule of law inherent in the first case is announced; then the rule of law is made applicable to the second case. This is a method of reasoning necessary for the law, but it has characteristics which under other circumstances might be considered imperfections. (pp. 1-2)

Analogical reasoning can be successfully trained. Feuerstein, Schlesinger, Shalom, and Narrol (1972) and Feuerstein (1979) have presented the results of an extensive training program for verbal and figural analogies that constitute part of Feuerstein's Learning Potential Assessment Device (LPAD). Feuerstein and his colleagues have used two basic kinds of training, which might be termed "performance-componential" and "metacomponential." In an experiment involving 551 children (including urban, upper-middle class school children; urban, lower-class school children; and educable mentally retarded children), Feuerstein (1979) found significant effects of verbal and figural training of both kinds. The largest gains were made, however, by children receiving both kinds of training together.

Whitely and Dawis (1973) described a "cognitive intervention for improving the estimate of latent ability measured from analogy items," and Whitely and Dawis (1974) tested this and other interventions on high school students. There were six conditions varying amount of practice, instruction,

and feedback. The "practice" groups did not perform better than the controls, but the "instruction" groups did perform better. Experimental groups receiving feedback did not perform better than groups not receiving it, but the group receiving semantic category instruction in addition to feedback performed significantly better than the comparable group not receiving category instruction. These and other results suggested that the category instruction was critical to improved performance.

Sternberg and Ketron (1980) have found that it is possible to train many subjects to use the variants of the models described earlier. Subjects are shown how to use the strategies to solve analogies composed of schematic pictures, and left to their own devices, the subjects can continue to use these strategies. Instruction was successful for analogies with integral attributes, but not for analogies with separable attributes.

Are certain strategies better for some people and other strategies better for others? There is some evidence that is at least suggestive that this is the case. Sternberg and Rifkin (1979) and Sternberg and Nigro (1980) have found that older subjects tend to be more nearly exhaustive in their information processing. More nearly exhaustive strategies tend to increase accuracy, but at the expense of requiring a greater memory load during analogy solution. The data suggest that subjects should be encouraged to use a strategy that is maximally exhaustive but that does not exceed the capacity of their working memory in its demands. With developments in technology for measuring working memory capacity (see Case, 1978a, 1978b), it may be possible eventually to individualize instruction in analogy solution in a way that maximizes individual analogy performance.

To recapitulate, the proposed metatheoretical framework can be and has been applied to the understanding of one form of inductive reasoning, reasoning by analogy. We now turn to a consideration of how the framework can be applied to the study and understanding of deductive reasoning.

Deductive Reasoning

The Scope of Deductive Reasoning

In deductive reasoning, the information contained in the premises of a problem is logically (although not necessarily psychologically) sufficient to reach a valid conclusion. A number of different kinds of deductive-reasoning problems have been studied, among them:

1. Linear syllogisms, e.g., "John is taller than Pete; Pete is taller than Bill; who is tallest?" In these problems, a logically valid conclusion is implied by the premises only if it is assumed that the relations linking the terms are transitive. For example, the relation "taller than" would satisfy transitivity, whereas the relation "plays better tennis than" might not. The problems may be presented in either verbal or nonverbal form, and in either case, may be embellished by the addition of negatives (e.g., "John is not as tall as Pete") or even additional premises. When premises are added, the problems are usually referred to as linear ordering or transitive inference problems, and indeed, linear syllogisms may be viewed, strictly speaking, as one of many possible kinds of linear ordering problem. Reviews of the literature on linear syllogistic reasoning can be found in Johnson-Laird (1972) and in Sternberg (1980b), as well as below. Some of the original sources include Clark (1969a, 1969b, 1971, 1972); DeSoto, London, and Handel

(1965); Donaldson (1963); Handel, DeSoto, and London (1968); Hunter (1957); Huttenlocher (1968); Huttenlocher and Higgins (1971, 1972); Huttenlocher, Higgins, Milligan, and Kauffman (1970); Huttenlocher and Strauss (1968); Keating and Caramazza (1975); Lutkus and Trabasso (1974); Piaget (1921, 1928, 1955, 1970); Potts (1972, 1974); Potts and Scholz (1975); Riley and Trabasso (1974); Shaver, Pierson, and Lang (1974); Sternberg (1980a, 1980b, 1980c); Trabasso (1975); Trabasso and Riley (1975); Trabasso, Riley, and Wilson (1975); and Wood, Shotter, and Godden (1974).

2. Categorical syllogisms, e.g., "All dorfles are dingbats. Some dunkits are dorfles. Can one conclude that some dunkits are dingbats?" Premises of categorical syllogisms, like those of linear syllogisms, can be presented in either affirmative or negative form (e.g., "No dorfles are dingbats" or "Some dunkits are not dorfles"). Premises of categorical syllogisms, unlike those of linear syllogisms, however, are almost never presented in pictorial form, and as a result, the syllogisms have not been presented to very young (preoperational) children. The syllogisms may be presented with more than two premises, in which case they are called "sorites," or "set inclusion problems." Reviews of the literature on categorical syllogistic reasoning can be found in Guyote and Sternberg (1978) and in Wason and Johnson-Laird (1972). Some of the original sources include Begg and Denny (1969); Ceraso and Provitera (1971); Chapman and Chapman (1959); Dickstein (1975, 1976, 1978); Erickson (1974, 1978); Frase (1966a, 1966b, 1968); Gilson and Abelson (1965); Griggs (1976, 1978); Henle (1962); Henle and Michael (1956); Janis and Frick (1943); Johnson-Laird (1975); Johnson-Laird and Steedman (1978); Kaufman and Goldstein (1967); Lefford (1946); Lippman

(1972); McGuire (1960); Morgan and Morton (1944); Revlin and Leirer (1978); Revlis (1975a, 1975b); Richter (1957); Roberge and Paulus (1971); Sells (1936); Simpson and Johnson (1966); Sternberg and Turner (in press); Wason and Johnson-Laird (1972); Wilkins (1928); Wilson (1965); and Woodworth and Sells (1935).

3. Conditional syllogisms, e.g., "If Conrad the Clown performs, people laugh; Conrad the Clown performs; can one conclude that people laugh?" As is the case with categorical syllogisms, premises may be negated (e.g., "If Conrad the Clown performs, people do not laugh," or "Conrad the Clown does not perform"). The premises may also be strung together to form an arbitrary number of items. Problems are almost always presented verbally. Reviews of the literature can be found in Guyote and Sternberg (1978) and in Wason and Johnson-Laird (1972). Some of the original sources include Guyote and Sternberg (1978); Kodroff and Roberge (1975); Marcus and Rips (1979); Osherson (1974, 1975); Paris (1973); Rips and Marcus (1977); Roberge and Paulus (1971); Staudenmayer (1975); Staudenmayer and Bourne (1977); Taplin (1971); Taplin and Staudenmayer (1973); Taplin, Staudenmayer, and Taddonio (1974); Wason and Johnson-Laird (1972).

These three kinds of syllogistic reasoning are not, of course, the only kinds of deductive reasoning that have been or might be studied, but they account for a fairly large proportion of the literature on deduction. Other kinds of deduction problems are considered by Wason and Johnson-Laird (1972), and a good logic text such as Copi (1978) reviews a wide range of types of deduction problems. Rather than attempting to survey this entire literature here within the framework proposed in the preceding section, we

shall review just one literature--that on linear syllogistic reasoning--as an example of the form such a review takes. Space considerations and the scope of this chapter simply do not permit a review of the entire range of literature on deductive reasoning.

A Case Study of Deductive Reasoning: The Linear Syllogism

1. Nature of the problem. In a linear syllogism, an individual is presented with two premises, each describing a relation between two items, one of which overlaps between premises. The individual's task is to use the overlap to determine the relation between the two items not occurring in the same premise, and then to answer a question requiring knowledge of this relation. In the item, "Sue is older than Lil; Lil is older than Ann; who is oldest?" the individual must deduce that Sue is older than Ann, and hence, that Sue is oldest of the three girls.

The standard domain of linear syllogisms consists of 2^5, or 32 item types, gotten by allowing each of the three adjectives in the problem (one in each premise and one in the conclusion) to be at either one pole (e.g., old) or at the other pole (e.g., young); by allowing the premises to be either affirmative (as above) or negative equative (e.g., "Lil is not as old as Sue"); and by allowing the correct answer to the item to be in either the first premise (as above) or in the second premise (as would be gotten by reversing the order of the premises). The premise adjectives are usually not psychologically symmetrical. One form (in this case, old) is simpler in certain senses to be described than the other form (in this case, young). The simpler form is the one that constitutes the name of the scale (in this

case, oldness rather than youngness, or in a similar case, tallness rather than shortness). The simpler form is referred to as unmarked, whereas the more complex form is referred to as marked.

Although the 32 items formed in the above way constitute the standard domain of linear syllogisms, they are not the only possible linear syllogisms. Additional items may be formed by allowing just one or the other premise to be negated, or by allowing problems to be indeterminate, i.e., specifying only a partial rather than a full ordering, e.g., "Sue is older than Lil; Sue is older than Ann; who is oldest?" The theories to be described below need augmentation in order to deal with the additional complexities created by items of these kinds.

Performance on linear syllogisms satisfies the four criteria for a "worthwhile" measure described in the preceding section. First, performance can be quantified by measurement of either response latency or error rate. Second, it can be measured reliably: Reliabilities of latencies (across item types) are generally in the high .80's or low .90's (Sternberg, 1980c, 1980d). Third, construct validity has been demonstrated numerous times in various ways: The linear syllogism plays an important part in Piaget's (1921, 1928, 1955) developmental theory of intelligence, since the ability to perform transitive inferences is alleged to differentiate preoperational from concrete-operational children; the problem plays an important role in DeSoto's theory of people's predilections for linear orderings (DeSoto, 1961; DeSoto, London, & Handel, 1965), in that the problem permits formation of a single, linear ordering; the problem plays an equally important role in Clark's (1969b, 1973) theory of linguistic processes in verbal comprehension, in that the processes alleged to be used

in solving linear syllogisms are alleged also to be used in a large variety of verbal-comprehension tasks; and the problem plays a central role in my own unified componential theory of human reasoning(Sternberg, 1978a, 1979b), in that the processes used are alleged to be used in a variety of deduction tasks, and the task itself falls into the task hierarchy that constitutes the organization of the theory. Finally, the empirical validity of individuals' performance on the problem has been demonstrated repeatedly. Burt (1912) used the problem in measuring the intelligence of school children, and performance on the problem has been found to be highly correlated with performance on verbal, spatial, and abstract reasoning ability tests (Shaver, Pierson, & Lang, 1974; Sternberg, 1980c; Sternberg & Weil, 1980): Correlations with such tests generally fall in the range from .30 to .60.

2. Performance components. Theorists differ in their views regarding the performance components used in solving linear syllogisms. Three views will be discussed here, based upon three different models of linear syllogistic reasoning. The three models are a spatial model based upon the DeSoto et al. (1965) and Huttenlocher (Huttenlocher, 1968; Huttenlocher & Higgins, 1971) models, a linguistic model based upon the Clark (1969b) model, and a linguistic-spatial mixed model described by Sternberg (1980d). Although the first two information-processing models are based upon previous models, they are not isomorphic to these previous models, in that the first two models were not presented by their original formulators in "componential" terms. The present models do seem to capture the major intuitions of the models as originally proposed.

The models all agree that there are certain encoding, negation, marking, and response operations that contribute to the latency with which a subject solves a linear syllogism. All full linear syllogisms contain

certain terms and relations to be encoded, and require a response. Only some linear syllogisms contain premises with negations and marked adjectives. Although the models agree on the presence of these performance components, they disagree as to which of these components are spatial and which are linguistic. The three models of linear syllogistic reasoning will be presented with reference to an example of a relatively difficult linear syllogism: _C_ is not as tall as _B_; _A_ is not as short as _B_; who is shortest? The correct answer is _C_, and by convention, _A_ will always refer to the extreme item at the unmarked end of the continuum, and _C_ to the extreme item at the marked end of the continuum. Each of these models can be represented in flow-chart form, and detailed descriptions of the various models are presented elsewhere (Sternberg, 1980d). Johnson-Laird (1972) has proposed slightly different flow charts for two of the models, the spatial and the linguistic ones. I describe here in detail only the mixed model, which is best supported by the available data (Sternberg, 1980c, 1980d).

In the linguistic-spatial _mixed model_, linguistic decoding of the problem is followed by its spatial recoding. The subject begins solution by reading the first premise. In order for the premise to be understood, it must be formulated in terms of the kind of deep-structural propositions proposed by the linguistic model. Encoding a marked adjective into this deep-structural format takes longer than encoding an unmarked one. Also, the presence of a negation requires a reformulation of the deep-structural proposition. Thus, "_C_ is not as tall as _B_" is originally formulated as (_C_ is tall+; _B_ is tall), and is then reformulated as (_B_ is tall+; _C_ is tall), as in the linguistic model. Once the deep-structural propositions for the premise are in final linguistic form, the terms of the propositions are

seriated spatially. If there is a marked adjective, the subject takes additional time in seriating the relation spatially in the nonpreferred (usually bottom-up) direction. If the adjective is not marked, then the premise is seriated in the preferred (usually top-down) direction. Note that whereas a negation is processed linguistically, a marked adjective is processed first linguistically (in comprehension) and later spatially (in seriation). After seriating the first premise, the subject repeats the steps described above for the second premise.

In order for the subject to combine the terms of the premises into a single spatial array, the subject needs the pivot available. The pivot is either immediately available from the linguistic encoding of the premises, or else it must be found spatially. According to the mixed model, there are two ways in which the pivot can become available immediately: (a) It is the single repeated term from all previous linguistic encodings; or (b) it is the last term to have been linguistically encoded. These rules have different implications for affirmative and negative premises.

In problems with two affirmative premises, the pivot is always immediately available, since each premise has been linguistically encoded just once. One term, the pivot, is distinctive from the others in that more than one relational tag has been associated with it, one from its encoding in the first premise, and one from its encoding in the second premise. The other two terms each have just a single relational tag associated with them. The second principle therefore need not even be applied. Indeed, it is applied only if the first principle fails.

The use of distinctiveness as a cue to the identity of the pivot fails in problems with at least one negative premise. In these problems, each premise

containing a negation is encoded in two different ways--in its original encoding and in its reformulated encoding in which the roles of their terms have been reversed. The pivot is therefore no longer the only term with more than one relational tag associated with it, and it thus loses its distinctiveness. The subject must therefore search for the term with the largest number of relational tags, unless he or she can apply the second principle.

When the distinctiveness principle fails, the subject attempts to link the first premise to the last term to have been encoded in working memory. If this term of the second premise happens to be the pivot, the link is successful, and the subject can proceed with problem solution. Pivot search can thus be avoided if the last term to have been encoded is the pivot. But if this term is not the pivot, the link cannot be made, and the subject must search for the pivot--the term with the largest number of relational tags. This search for the pivot takes additional time.

Once the pivot has been located, the subject seriates the terms from the two spatial arrays into a single spatial array. In forming the array, the subject starts with the terms of the first premise, and ends with those of the second premise. The subject's mental location after seriation, therefore, is in that half of the array described by the second premise. The subject next reads the question. If there is a marked adjective in the question, the subject will take longer to encode the adjective, and to seek the response at the nonpreferred (usually bottom) end of the array. The response may or may not be immediately available. If the correct answer is in the half of the array where the subject just completed seriation (his or her active location in the array), then the response will be available immediately. If the question requires an answer from the other half of

the array, however, the subject will have to search for the response, mentally traversing the array from one half to the other and thereby consuming additional time.

One final search operation is used optionally under special circumstances. If the subject has constructed a sharp spatial encoding, then he or she is now ready to respond with the correct answer. If the subject's encoding is fuzzy, however, the subject may find that he or she is unable to respond with a reasonable degree of certainty. The subject therefore checks his or her tentative response as determined by the spatial representation with the encoding of that response term in the linguistic representation. If the question and response are congruent, the check is successful, and the subject reformulates the question to ascertain whether it can be made congruent with the response. Only then does he or she respond.

To summarize, the performance components needed for solution of linear syllogisms according to the linguistic-spatial mixed model are premise reading, encoding of terms in the preferred relation, encoding of terms in the nonpreferred relation (which may be viewed as an augmentation of the preceding component), reversal of terms in the encoded relation for negated premises, seriation of terms in the preferred direction, seriation of terms in the nonpreferred direction (which may be viewed as an augmentation of the preceding component), pivot search, seriation of terms for the combined premises, question reading, response search, establishment of congruence between question and response (optional), and response.

All of the models are rather detailed, and it is easy to lose the forest among the trees. Thus, some overall comparison among the three models may help put them into perspective. The models all agree that marked adjectives and negations should increase solution latency. They disagree, however, as

to why solution latency is increased. According to the spatial model, solution latency is increased because processing of negations and marked adjectives requires a more complex encoding of information into a visualized spatial array. According to the linguistic model, the additional time results from increased difficulty in a linguistic encoding process. According to the mixed model, negations require a more complex linguistic encoding process, whereas marked adjectives require first more complex linguistic encoding and then more complex spatial encoding.

The models also agree that some form of pivot search (for the middle term in the array) is needed under special circumstances. The models disagree, however, as to what these circumstances are. In the spatial model, pivot search is required for premises that are not end-anchored, that is, for premises in which the first term is the middle rather than an end of a spatial array. Absence of end-anchoring necessitates a search through the visualized spatial array. In the linguistic model, pivot search results from compression of the first premise in the deep-structural encoding (i.e., the first term of the first premise, but not the second, is stored in working memory). If the term that was dropped from working memory in compression happens to be the pivot term, then the subject has to retrieve that term back from long-term memory. In the mixed model, pivot search is required if the reformulated deep-structural version of a negative second premise does not have the pivot in its latter (and hence most recently available) proposition.

The spatial and mixed models agree that the terms of the two premises are combined into a single, unified representation. This combination is accomplished through a seriation operation in which each of two partial

spatial arrays is unified into a single array. The linguistic model disagrees: Functional relations from the two premises are stored separately.

The linguistic and mixed models agree in the need for an operation to establish congruence between question and answer, but in the mixed model, the establishment of congruence is optional. It is used only when the spatial encoding of terms is of insufficient quality to permit the subject to respond to the problem with a reasonable degree of certainty. No operation for the establishment of congruence exists in the spatial model.

In the spatial model, subjects are hypothesized to prefer working in a certain direction (usually top-down) between as well as within premises. Generally, this preference means that extra time will be spent in seriation if the term at the preferred end of the array does not occur in the first premise. No corresponding "additional latency" exists in either the linguistic or mixed model.

In the linguistic model, subjects search the deep-structural propositions for the term that answers the question. In a spatial array, it is obvious which term corresponds to which question adjective. For example, the tallest term might be at the top, the shortest term at the bottom. In linguistic propositions, there is no such obvious correspondence, so that the subject must check both extreme terms relative to the pivot, seeking the correct answer.

In the mixed model, subjects have to search for the response to the problem if their active location in their final spatial array is not in the half of the array containing the response. Subjects mentally traverse the array to the other half, looking for the response. No corresponding operation exists in either the spatial or linguistic model.

Finally, the models agree that the final operation is a response process, whereby the subject selects his or her answer.

These are not the only models of linear syllogistic reasoning that have been or might be proposed. Hunter (1957) and Quinton and Fellows (1975), for example, have presented alternative models that can also be cast in "componential" terms. The three models presented seem to be the major models of current interest, however, and are probably the ones worthy of the closest attention, at least at the present time.

In order to compare the relative abilities of the models to account for performance of human subjects on linear syllogisms, it is necessary first to postulate a combination rule. Discussion of the relative merits of the models will therefore be postponed until combination rules are discussed below.

3. Representation of information. Theorists disagree as to the form of representation subjects use for information stored, manipulated, and retrieved in the course of linear syllogistic reasoning. The basic controversy has been over whether information is represented spatially, linguistically, or both spatially and linguistically. Spatial theorists argue that information is represented in the form of a spatial array that functions as an internal analogue to a physically realized or realizable array. Linguistic theorists argue that information is represented in the form of linguistic, deep-structure propositions of the type originally proposed by Chomsky (1965). Mixed model theorists argue that both forms of representation are used, with the linguistic form of representation serving its primary function during initial decoding of the problem, and with the spatial form of representation serving its primary function during later recoding of the problem. A reso-

lution of this controversy would not only enlighten us with regard to transitive inference, but might further shed light on the kinds of arguments that are valuable in distinguishing between subjects' use of spatial or imagerial representations on the one hand, and linguistic or propositional representations on the other (see Anderson, 1976; Kosslyn & Pomerantz, 1977; Pylyshyn, 1973).

Let us first consider evidence in favor of a spatial representation. Eight principal kinds of evidence have been adduced in favor of a spatial representation for information.

i. Introspective reports. Many subjects in various experiments have reported using spatial imagery to solve transitive inference problems such as linear syllogisms (Clark, 1969a; DeSoto et al., 1965; Huttenlocher & Higgins, 1971).

ii. Need for spatial array to combine premise information. At some point during the course of problem solution, subjects must comprehend the higher-order relation between the two lower-order relations expressed in the individual premises. Such comprehension is tantamount to making the transitive inference needed to solve the problem. Spatial imagery theorists have specified at a reasonable level of detail how such comprehension can take place (see Huttenlocher, 1968; Huttenlocher & Higgins, 1971). Linguistic theorists, however, have not specified in reasonable detail how the transitive inference is actually made. Clark (1971) has admitted that the "linguistic theory is not complete. For one thing, it does not fully specify how information from the two premises are [sic] combined" (p. 513).

iii. Comparability of data patterns for purported imaginal arrays to those for physical arrays. One of Huttenlocher's main arguments in favor of spatial imagery has been that "the difficulty of solving different forms of [linear] syllogisms parallels the difficulty of arranging real objects according to comparable instructions" (Huttenlocher et al., 1970). A series of experiments has shown that the two types of items do indeed show parallel patterns of data (Huttenlocher, Eisenberg, & Strauss, 1968; Huttenlocher et al., 1970; Huttenlocher & Strauss, 1968).

iv. Symbolic distance effects. Data reported by Potts (1972, 1974) and by Trabasso and his colleagues (Trabasso & Riley, 1975; Trabasso, Riley, & Wilson, 1975) seem strongly to implicate some kind of spatial process in linear ordering problems. In a typical experiment, subjects are taught a linear ordering of items that takes the form (A, B, C, D, E, F). Subjects are trained only on adjacent pairs of items. Subjects are able to judge the untrained relation between B and E more rapidly than they are able to judge the trained relation between C and D. The further apart the two items are, the easier the judgment turns out to be. This symbolic distance effect is compatible with the kind of "internal psychophysics" proposed by Moyer (1973) and by Moyer and Bayer (1976), whereby a spatial analogue representation is constructed for the array, and elements of this analogue representation are compared to one another.

v. Serial position effects. In the linear-ordering experiments described above, subjects are trained on all adjacent pairs of items in the linear ordering. Trabasso and his colleagues (Lutkus & Trabasso, 1974;

Riley & Trabasso, 1974; Trabasso et al., 1975) have found that errors made during training and retraining exhibit a serial-position effect with respect to position of the pairs in the linear ordering: Maximum errors occur on middle pairs, and fewer errors occur on pairs nearer the ends of the ordering. This serial-position effect is interpreted as prima facie evidence for an underlying spatial array (see Bower, 1971).

vi. Directional preferences within linear orderings. In many of the adjective pairs used in linear syllogism problems, one adjective of a bipolar pair results in more rapid or more accurate solution than the other. For example, use of the adjectives taller and better results in facilitated performance relative to the adjectives shorter and worse (Handel et al., 1968). These authors have proposed that faster solution for the adjectives taller and better can be accounted for by the facts that (a) taller-shorter is represented along a continuum proceeding from top to bottom and better-worse is represented along a continuum proceeding from right to left, and (b) people proceed more readily in a downward direction than in an upward direction, and in a rightward than in a leftward direction (p.513).

vii. End-anchoring effects. Investigators of transitive inference have repeatedly found end-anchoring effects in their data (see DeSota et al., 1965; Huttenlocher, 1968). End-anchoring effects are observed when it is easier to solve a transitive inference problem presented from the ends of an array inward than it is to solve the problem presented from the middle of the array outward. Such effects are consistent with a spatial representation of information.

viii. Correlations with spatial visualization tests. Shaver, Pierson, and Lang (1974) have reported correlations across subjects between errors in the solution of linear syllogisms and scores on tests of spatial visualization.

These correlations varied in magnitude, but an impressive number of them reached statistical significance. These correlations were interpreted as evidence that spatial imagery is used in the solution of linear syllogisms.

With eight kinds of evidence converging on the same conclusion, one is tempted to accept the conclusion without further ado. Yet, none of the eight kinds of evidence proves to be conclusive considered either by itself or in conjunction with the remaining kinds of evidence.

Consider first introspective reports. Introspective reports of the use of imagery are common, and are acknowledged even by the most prominent linguistic theorist (Clark, 1969b). A long-standing question in psychology, however, has been whether such reports can be accepted at face value (see, for example, Nisbett & Wilson, 1977). Although such reports are certainly suggestive, they are certainly not conclusive. Consider next combination of premise information, symbolic distance effects, serial position effects, and end-anchoring effects. Can a linguistic representation account for any or all of these effects? The answer appears to be affirmative: A small modification and extension of a linguistic representation suggested by Holyoak (1976) will predict all of these effects. Consider next comparability of data patterns for imaginal and physical arrays. Huttenlocher's argument that data patterns for reasoning with purported imaginal arrays are very similar to those for placement with actual physical arrays presents a reasonable case for the analogy between the two kinds of arrays. The correspondence does not always hold, however (Clark, 1969b, 1972a). Consider now directional preferences. In general, adjectives that encourage top-down or right-left processing are also those that are linguistically unmarked. Thus, linguistic theory also predicts facilitated processing for these adjectives. Consider finally correlations with spatial tests. Available correlational evidence

from the Shaver et al. (1974) study provides convergent validation for the spatial hypothesis, but does not provide discriminant validation with respect to one or more alternative hypotheses. In other words, errors on the linear syllogisms task might well have correlated with tests of spatial visualization ability because of a general factor that pervades performance on both spatial and linguistic ability tests. In order to provide a stronger test of the spatial hypothesis, one would have to show high correlations between linear syllogism and spatial test performance coupled with low correlations between linear syllogism performance and linguistic test performance.

Consider now evidence favoring a linguistic representation of information in linear syllogistic reasoning. Three principal kinds of evidence have been adduced in its favor.

The first, the principle of primacy of functional relations, states that "functional relations, like those of subject, verb, and direct object, are stored, immediately after comprehension, in a more readily available form than other kinds of information, like that of theme" (Clark, 1969b, p. 388). This principle forms the basis for the linguistic representation of information in terms of base strings and underlying deep-structural transformations on these base strings. Clark has not offered any direct experimental evidence to support the principle, although he does claim indirect support from several sources (Donaldson, 1963; Piaget, 1928).

The second kind of evidence is the principle of lexical marking. According to Clark's (1969b) lexical marking principle, "the senses of certain 'positive' adjectives, like good and long, are stored in memory in a less complex form than the senses of their opposites" (p. 389). The "positive" adjectives are the unmarked ones, and their opposites are the marked ones. If, as Clark claims, marked adjectives are stored in memory in a more linguistically complex form than is needed for unmarked adjectives, then one might well

expect the encoding of marked adjectives to be more time-consuming than the encoding of unmarked adjectives, and indeed, all studies of linear syllogistic reasoning that have investigated both marked and unmarked adjectives have found longer latencies or more errors associated with items containing marked adjectives than have been found with items containing unmarked adjectives. This evidence therefore seems on its face to support the principle of lexical marking.

The third kind of evidence is the principle of congruence. According to Clark (1969b), "information cannot be retrieved from a sentence unless it is congruent in its functional relations with the information that is being sought" (p. 392). If information from the premises is not congruent with the information being sought, then additional time will be needed to establish congruence between the question and response. Suppose, for example, the question is "Who is best?" and the answer is *A*. If *A* were encoded from a premise such as "*A* is better than *B*," then solution should be relatively rapid, since *A* was encoded in terms of the comparative *better* and the question asks who is *best*. Suppose that instead, the relevant premise were "*B* is worse than A," which, according to Clark, can be expanded to "*B* is worse than *A* is bad." This premise does not contain information congruent with the question. The question can be answered only if it is reformulated to read, "Who is least bad?"

Evidence in favor of a linguistic representation of information is at least as flimsy as that in favor of a spatial representation. First, the observational evidence to support the principle of primacy of functional relations is suggestive at best, and certainly no stronger than subjects' direct introspective reports of spatial imagery. At present, the principle seems to stand more as a presupposition for the remaining two principles than as a principle that is testable in its own right. Second, the mere existence of a marking

effect as predicted by the _principle of lexical marking_ does not in itself argue for a linguistic representation for information. As noted earlier, a number of investigators have noticed that the unmarked form of a bipolar adjective pair is generally the form that would be expected to appear at the top of a spatial array (DeSoto et al., 1965; Huttenlocher & Higgins, 1971). If an adjective pair could be found in which the marked form suggested the top of a spatial array and the unmarked form suggested the bottom of a spatial array, then, according to Clark (1969b), it would be possible to disentangle the spatial and linguistic accounts of the marking effect. Such an adjective pair is found in _deep-shallow_, where _deep_, the unmarked adjective in the pair, suggests the lower end of a spatial array. Clark (1969b) has reported that when subjects are presented with linear syllogisms containing the adjective pair, _deep-shallow_, the standard marking effect is obtained. Another critical adjective pair, _early-late_, is reported by Clark (1969b), however, to show results opposite to those predicted by lexical marking. Finally, consider again the _principle of congruence_. Spatial theorists are skeptical that the available data provide adequate support for the principle of congruence. In a series of recent experiments, Potts and Scholz (1975) obtained a congruence effect under some circumstances but not under others. Clark's (1969b) data provide only weak support for the principle of congruence. My own data (Sternberg, 1980c) suggest that the "principle of congruence" holds when items are presented in standard form, but not when they are presented premise-by-premise, with subjects pacing the rate of premise presentation. I believe the reason for the difference can be found in the relative quality of encodings in the two kinds of experimental situations (see description of processes in mixed model presented earlier).

Finally, let us consider evidence in favor of a dual linguistic-spatial representation. In a series of studies (Sternberg, 1980c, 1980d; Sternberg & Weil, 1980), we have administered to subjects linear syllogisms plus psychometric tests of verbal and spatial abilities. The psychometric tests were factor analyzed in order to yield two orthogonal factors of measured verbal and spatial abilities. Overall response latencies and latencies for individual components of information processing (determined according to the linguistic-spatial mixed model) were then correlated with the factor scores. In every one of six experiments, overall latencies were significantly correlated with both verbal and spatial factors. The absolute and relative magnitudes of the correlations with the two factors differed across experiments, but were all in the .30 to .60 range. Moreover, latency parameters hypothesized to represent the durations of processes operating upon a linguistic data base generally correlated with verbal but not spatial ability; latency parameters hypothesized to represent the durations of processes operating upon a spatial data base generally correlated with spatial but not verbal ability; and confounded latency parameters that represented components operating upon both kinds of data bases generally correlated with both abilities. These results seem consistent with the notion that subjects use both linguistic and spatial representations in their solution of linear syllogisms.

To conclude, there is some evidence that subjects use a spatial representation, and some evidence that subjects use a linguistic representation. In each case, the evidence argues in favor of the use of one kind of representation, but not in opposition to the use of the other kind of representation. The entire body of evidence in favor of one or the other kind of representation is thus fully consistent with the use of both kinds of representation, and the results from the Sternberg studies seem to argue that subjects do indeed use both

linguistic and spatial representations in their solution of linear syllogisms. The Sternberg results provide further evidence regarding which performance components of the mixed model operate upon which kind of representation. In general, "linguistic" parameters show correlations with linguistic but not spatial ability tests, "spatial" parameters show the reverse pattern, and confounded parameters show correlations with both.

4. Combination rule. All components were assumed to be executed in the order specified in the earlier description of models. All of the models were tested with an additive combination rule, i.e., a rule assuming strictly serial information processing. In a series of six experiments comparing the three alternative models for untrained adult subjects (Sternberg, 1980c, 1980d; Sternberg & Weil, 1980), the mixed model was best in each case. Values of R^2 between predicted and observed latency data ranged from .74 to .88 for the mixed model, with a median of .83. The range for the spatial model was .57 to .66, with a median of .58. The range for the linguistic model was from .59 to .69, with a median of .62. Averaged across the four experiments of Sternberg (1980d), the root-mean-square deviations (RMSD) of observed from predicted values were 28 centiseconds for the mixed model, 55 centiseconds for the spatial model, and 52 centiseconds for the linguistic model. All parameters of the mixed model were statistically significant in each of the six experiments (except for negation in one experiment), although many parameters of the other models were nonsignificant across the various experiments. Overall, then, these data can be interpreted as providing rather strong support for the mixed model. Unfortunately, mine are the only data comparing the three models, since the mixed model has only very recently been proposed for the first time.

5. Durations, difficulties, and probabilities of component execution. Parameter estimates for the mixed model were relatively stable across experiments (providing further support for the tenability of the mixed model), although most of the estimates of latency could be reduced by giving subjects instructions that encouraged speedy solution (Sternberg, 1980c).

There turned out to be unexpected complexities in the modeling of error rates, the explanation of which would be beyond the scope of this chapter. Estimates of parameters for predicting error rates are presented in Sternberg (1980c). Response probabilities were not modeled, because under standard instructions telling subjects to respond as quickly as possible without making errors, error rates ran only about 1%.

6. Metacomponents. Consider the six metacomponents of reasoning and problem solving identified in the previous section, and what we know about each one.

i. Selection of performance components for task solution. Analyses of individual model fits indicate that about 70-75% of subjects spontaneously choose the components of the mixed model, about 10-15% spontaneously choose those of the linguistic model, and about 10-20% spontaneously choose those of the spatial model. Some subjects, of course, use none of these models. Individual differences in the strategy components subjects spontaneously choose to use in solving linear syllogisms do not correspond in a systematic way to ability patterns (Sternberg & Weil, 1980).

ii. Selection of representation(s). When correlations of latency scores with ability factor scores are analyzed for subjects using each of the various models, it is found that scores for subjects using the linguistic-model com-

ponents correlate with verbal ability scores but not with spatial ability scores; scores for subjects using the spatial-model components correlate with spatial ability scores but not with verbal ability scores; and scores for subjects using mixed-model components correlate with both verbal and spatial ability scores. Thus, subjects seem able to choose a representation of information compatible with their selection of performance components (Sternberg & Weil, 1980).

iii. Selection of strategy for combining components. The linear models that have been tested all assume serial processing for combining components. These models provide a good, although imperfect fit to the data. Since nonlinear (and hence nonserial) models have not been tested, we really don't know how many subjects are strictly serial, and how many use at least some parallel processing in their solution of the problems.

iv. Decision as to whether to maintain a strategy. Data to be described shortly suggest that subjects generally stick with the strategy they start with. It is interesting to note that when a change in strategy is induced through experimental instructions, subjects react in different ways as a function of their initial success with the strategy they are using. The probability of their adopting a trained strategy that differs from the mixed strategy most of them use appears to be inversely related to the subjects' prior success with the mixed strategy. In other words, subjects who find themselves performing effectively ("winning") with the mixed strategy are less likely to switch to the trained alternative strategy, despite instructions to do so. They seem to know that they have a winning strategy, and to decide on this basis to stick with it (Sternberg & Weil, 1980).

v. <u>Selection of a speed-accuracy tradeoff</u>. Subjects can be instructed to alter their speed-accuracy tradeoff function in order to increase rate of information processing at the expense of accuracy of information processing. In one experiment, speed-emphasis instructions reduced mean latency of response from the typical 7 to 7 1/2 seconds to a faster 6 seconds, at the cost of an increase in error rate from 1% to 7%. The distribution of model use was unaltered by the speed instructions, even though no explicit mention was made of the use of any one model or another. When subjects speed up, most of the increase in rate of response is isolated in encoding operations. Pivot search, response search, and response also show some decrease in latency. Negation, marking, and noncongruence are only minimally affected (Sternberg, 1980c).

vi. <u>Solution monitoring</u>. The extent to which subjects monitor their performance during linear syllogistic reasoning is unknown. In general, subjects are able to give only a very vague account of how they went about solving the problems. Most subjects can report on whether or not they used imagery, but not much more. Their inability to describe their solution processes does not mean they do not monitor their performance, however, since solution monitoring can be conducted below the level of consciousness.

7. <u>Effects of problem format, problem content, and practice</u>. Each of these variables has received at least some study, so that we are in a position to assert at least tentatively some effects of these variables on performance in a linear syllogistic reasoning task.

Consider first <u>problem format</u>. Two basic procedures have been used in timing of performance. The most common procedure is to present the problems

for as long as subjects take to solve them. The subject's response to a problem terminates presentation of the problem (Clark, 1969b; Hunter, 1957; Huttenlocher, 1968; Sternberg, 1980a, 1980b, 1980c). An alternative and less common procedure is to present the linear syllogism for a period of 10 seconds. If a subject is able to solve the problem correctly in this amount of time, his or her response is counted as correct; otherwise, it is counted as an error (Clark, 1969a; DeSoto et al., 1965; Keating & Caramazza, 1975). The difference in procedure has a major effect upon inferences about subjects' processing strategies (Sternberg, 1980b). The former procedure tends to favor interpretation of results in terms of the mixed model, whereas the latter procedure tends to favor interpretation of results in terms of the linguistic model. If, however, the time limit at the deadline is changed, the results may favor the mixed model. Thus, the deadline determines what the results look like, for reasons discussed elsewhere (Sternberg, 1980c).

Two basic procedures have also been used with regard to presentation of the premises and question. Some investigators have presented both premises plus the question simultaneously (e.g., Clark, 1969a, 1969b; Hunter, 1957). Other investigators have presented the premises separately, or else have presented the premises together, but the question separately (e.g., Huttenlocher, 1968; Potts & Scholz, 1975). Still other investigators have used both procedures (Sternberg, 1980d), and even presented the question before the premises. The mixed model was always best. It seems that establishment of congruence is required only for the former procedure, where premise encoding tends to be less thorough and thus more in need of later verification (Sternberg, 1980d).

Consider next <u>problem content</u>. The effects of relational terms (usually adjectives) have been most thoroughly investigated by DeSoto et al. (1965) and Handel et al. (1968). Two characteristics of the relational terms have received the most attention: differences in directional preference between and within bipolar pairs, and differences in difficulty between and within bipolar pairs.

With regard to directional preferences, the research of DeSoto et al. and of Handel et al. has suggested that subjects tend to order certain relational pairs, such as <u>better-worse</u>, <u>father-son</u>, and <u>more-less</u>, vertically in spatial arrays. <u>Better</u>, <u>father</u>, and <u>more</u> are generally represented at the upper end of each array. Other relational pairs, such as <u>earlier-later</u> and <u>faster-slower</u>, tended to evoke horizontal spatial arrays, with <u>earlier</u> and <u>slower</u> at the left end of each array. In still other relational pairs, such as <u>cause-effect</u>, <u>farther-nearer</u>, and <u>lighter-darker</u>, most subjects are inconsistent in their directional preferences.

With regard to directional difficulties, Handel et al. (1968) tested subjects with problems containing a number of different relational pairs. Although these investigators did not explicitly test differences in item difficulty as a function of spatial direction, it is clear from their data that relational terms for which subjects were inconsistent in their spatial directions were more difficult to process than were relational terms for which subjects were consistent. Within relational pairs, DeSoto et al. (1965) and others have found that items are easier when presented with the adjective of a pair that encourages top-down rather than bottom-up processing, or left-right rather than right-left processing.

Consider finally effects of practice. Most theorists seem to assume that subjects are constant in their strategy. Not all theorists make this assumption, however. Citing the theory and data of Wood (1969), Wason and Johnson-Laird (1972) have proposed that

> the inexperienced subject represents the premises in a unified form (with or without imagery) because this is likely to be the normal practical mode of dealing with the relational information. But by dint of sheer repetition this approach is likely to give way to a purer and more formal strategy geared to the specific constraints of the problem. . . . In short, subjects seem likely to pass from an approach analogous to the IMAGE theory to one analogous to the LINGUISTIC theory. (p. 122)

According to this hypothesis, one would expect subjects to follow a spatial model early during their experience with linear syllogisms, and to switch later to a linguistic model.

Shaver et al. (1974) have proposed a strategy change hypothesis that reverses the sequence described above. They noted that Johnson-Laird (1972)

> hypothesized that imagery is abandoned in favor of a linguistic strategy after practice with three-term series problems. The opposite temporal sequence is indicated by our results, suggesting that in this case at least, imagery provided the "more economical and specialized" strategy. (p. 373)

According to this hypothesis, then, subjects are assumed to follow a spatial strategy early during their experience with linear syllogisms, and to switch later to a linguistic strategy.

Sternberg (1980d) tested these strategy change hypotheses in two ways. The first was to compare fits of the various models for earlier sessions of practice versus later sessions of practice. It was found that the mixed model was superior to the alternative models, without regard to session of practice, and that it was superior by roughly the same amount in each case. The second way of testing the hypothesis was to compare correlations of latency scores with verbal and spatial scores for early versus late sessions of practice. If a strategy change were occurring, one might expect the magnitudes of the correlations with spatial (or verbal) ability to decrease over sessions, and those of the correlations with verbal (or spatial) ability to increase over sessions. In fact, the relative magnitudes of the correlations remained about the same over sessions, again providing no evidence consistent with a strategy shift.

8. Individual differences within age level (adults). The general componential framework we have been using reveals a number of sources of individual differences. First, subjects differ in the components they use for solving linear syllogisms. The large majority of subjects appear to use the components of the mixed model, but nontrivial numbers use the components of either the spatial or linguistic models. These models, it must be remembered, are only approximations to subjects' actual strategies. The data of virtually all the subjects have reliable variance not accounted for by any of the models (Sternberg, 1980d). Second, subjects differ in their representations of information: Some appear to use only a linguistic

representation, others only a spatial representation; most appear to use both kinds of representation (Sternberg & Weil, 1980). Third, subjects appear to differ in the consistency with which they employ any strategy at all: The internal consistency reliability of individual data sets varies widely across subjects. Thus, subjects differ not only in the model which best fits their individual data sets, but in the extent to which any : can and does fit their data set at all. Fourth, subjects differ widely in the rates at which they execute the various performance components, with the largest individual differences occurring in the encoding and response operations (Sternberg, 1980d). Finally, subjects differ in their accuracy of component execution (Shaver et al., 1974).

9. Differences across age levels. Investigators have differed in their claims regarding what model children of various ages use in solving linear syllogisms or other kinds of transitive inference problems, but with the exception of Piaget (1921, 1928, 1955), they have been remarkably consistent in their claims that there is no evidence of strategy change across ages (Bryant & Trabasso, 1971; Hunter, 1957; Keating & Caramazza, 1975; Riley, 1976; Riley & Trabasso, 1974; Sternberg, 1980b; Trabasso, 1975). And even Piaget makes no claims of changes in strategies for children at or above the level of concrete operations. Sternberg's (1980b) data are the only ones that compare the spatial, linguistic, and mixed models across age levels, as well as the algorithmic model of Quinton and Fellows (1975). The mixed model outperforms the others at the grade 3, 5, 7, and 11 levels. At grade 9, there is an inversion, with the linguistic model outperforming the mixed model. This inversion, however, can be localized to the first

session of practice; in the second session, the mixed model performs better. Whether this represents a true phenomenon or a quirk in the data cannot be known, although this one inversion seems to fit into no particular pattern.

As one might expect, both solution latencies and error rates decrease with increasing age: Mean latencies (in seconds) for the 32 standard linear syllogisms are 14.51, 11.98, 10.02, 9.88, and 7.54 in grades 3, 5, 7, 9, and 11 respectively; mean error rates are .40, .25, .23, .18, and .16 for the respective grades. The largest decrease in latency across grade levels is in the response component. The other components also show generally decreasing trends, although the rates of decrease are much slower than that of the response component (Sternberg, 1980b).

10. Relationship between linear syllogistic reasoning and other kinds of reasoning. As would be expected, performance on linear syllogisms is significantly correlated with performance on other kinds of reasoning tasks.

Sternberg (1980d) reported a correlation of -.52 between latencies for linear syllogisms and scores on tests of abstract reasoning ability. (Negative correlations result from correlating latencies, where lower scores indicate superior performance, with test or factor scores, where higher scores indicate superior performance.) Although we have not correlated performance on linear and categorical syllogisms directly, the fact that both show high correlations with spatial ability tests would suggest that they would show high correlations with each other as well (Guyote & Sternberg, 1978; Sternberg, 1980d).

These high correlations can be explained at least in part within the present componential metatheoretical framework. First, linear syllogisms require at least some of the same performance components as do related kinds

of problems, such as categorical syllogisms: Both require encoding of premise information, decoding of negations, combination of information from pairs of premises, and response. Second, linear syllogisms are like categorical syllogisms and certain other kinds of reasoning problems in their requirement of a spatial representation of information for solution (at least for most people). Third, the problems share many of the same metacomponents with other kinds of reasoning problems. Regardless of whether or not the performance components are the same, in each case, the decision must be made as to what performance components to use, and similarly, decisions must be made regarding combination rule, representation, speed-accuracy tradeoff, etc. Thus, even if the content of the decisions differs, the acts of deciding are required in each case. Finally, the acquisition, retention, and transfer components used to learn, remember, and generalize performance on various kinds of reasoning tasks are probably highly overlapping, leading to increased correlations between tasks due to the similar psychological histories of the tasks.

11. Relationship between linear syllogistic reasoning and intelligence. Since all of the various kinds of reasoning tests mentioned above have been used in standard batteries for the assessment of intelligence, and since performance on linear syllogisms is rather highly correlated with performance on these other reasoning tasks, there is empirical evidence of the usefulness of linear syllogisms as psychometric measures of intelligence. Linear syllogisms have also played major roles in the two other major traditions of theory and research on intelligence, the Piagetian tradition and the information-processing tradition. In the Piagetian tradition, as mentioned earlier,

linear syllogisms have served as a basis for distinguishing preoperational children from concrete-operational ones. In the information-processing tradition, theorists have argued that the processes used in linear syllogistic reasoning are central to intelligent language comprehension (Clark, 1973), imaginal representation of linear orderings (DeSoto et al., 1965), and deductive reasoning in general (Sternberg, 1980d), which is an important aspect of intelligence.

12. Practical relevance. Virtually all of the theorists who have studied linear syllogisms have done so at least in part because of the practical importance of the processes underlying linear syllogistic reasoning, whether for everyday language comprehension, or whatever. Sternberg (1980d) has given an example of how transitive inference is used in many of the mundane situations of everyday life. Consider the plight of a customer eating at a restaurant. He or she is faced with what may be a bewildering choice of meals. Since the customer has neither the time nor the patience to compare very possible meal in order to determine which he or she prefers most, the customer relies upon a strategy of transitive inference, deciding that if, for example, pizza is preferred to an omelette, and an omelette is preferred to a garden salad, then pizza is preferred to the garden salad. Without transitive inference, every possible paired comparison would have to be done in order to be sure that one's preferred meal is being ordered. Consider another example, the task college admissions officers face in filling a small number of slots in the entering class from a large number of applications for those slots. Were the admissions officers not to make transitive inferences, the number of paired comparisons that would be

required to compare each candidate against each other candidate would be way beyond all reasonable bounds.

Sternberg and Weil (1980) were interested in whether particular strategies for solving linear syllogisms could be trained, and in whether an aptitude-strategy interaction exists in linear syllogistic reasoning whereby the efficacy of a particular strategy depends upon a person's pattern of verbal and spatial abilities. If the answers to both of these questions were affirmative, then it might be possible to train people to use a strategy that is optimal for their pattern of abilities. In fact, the answers to both questions were affirmative.

To recapitulate, the proposed metatheoretical framework can be and has been applied to the understanding of one form of deductive reasoning, linear syllogistic reasoning. Similar analyses have been performed for other forms of deductive reasoning (e.g., categorical and conditional syllogistic reasoning), but the purpose of this chapter is to show how the framework can be applied to a variety of problems, and so we will now turn to a consideration of rather different kinds of problems.

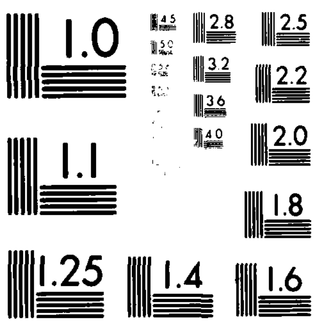
PROBLEM SOLVING AND INTELLIGENCE

Problem situations--the bases for problem-solving behavior--have been characterized in a number of different ways. Johnson (1955) has suggested that a problem situation exists when an individual's first goal-directed response is unrewarding. Köhler (1925) has suggested that a problem situa- tion exists when an individual must take an "Umweg," or detour, to reach a goal. Vinacke (1952) has taken a similar position, claiming that a situation exists when there is an "obstacle" to overcome. Woodworth and Schlosberg (1954) have argued that a problem situation exists when

AD-A086 796 YALE UNIV NEW HAVEN CT DEPT OF PSYCHOLOGY F/G 5/10
REASONING, PROBLEM SOLVING, AND INTELLIGENCE.(U)
APR 80 R J STERNBERG N00014-78-C-0025
UNCLASSIFIED RR-3-80 NL

2 OF 2
AD
A086796

08679



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

vidual has a goal, but no clear or well-learned route to the goal. Still other definitions have been proposed by Humphrey (1951), Maltzman (1955), Ray (1955), Russell (1956), Underwood (1952), and van de Geer (1957). According to Duncan (1959), who reviewed what was once "recent research on human problem solving," "the defining characteristics most frequently mentioned are the integration and organization of past experience when the definition refers to all of thinking, and the dimension of discovery of correct response when reference is made to problem solving specifically" (p. 397).

The definition I prefer is one offered by Raaheim (1974), which finds its historical roots in an earlier definition by Morgan (1941). Morgan suggested that a problem situation exists when there are some elements or conditions that are known, some other elements or conditions that are unknown, and when the solution depends upon a discovery of how to deal with the unknown factors of the situation. In Raaheim's words, a problem situation is a "deviant member of a series of earlier situations of the same sort" (p. 22). Thus, one always has some basis for dealing with the problem on the basis of past experience, but not enough of a basis to provide an immediate solution.

Problems may be subdivided in any of a number of ways. A convenient way of subdividing them, and the one that I will use for the present purpose, is in terms of "well-defined" and "ill-defined" problem spaces (see Newell & Simon, 1972). A problem with a well-defined problem space is one for which the steps to solution can be clearly specified by the experimenter, and, ultimately, by the problem solver. Problems of this kind often require a

series of small transformations on the problem input to yield the problem output. The particular difficulty is usually not in achieving any one step, but in achieving a coordinated set of steps that will yield the desired outcome. A problem with an ill-defined problem space is one for which the steps to solution cannot be clearly specified by either the experimenter or the problem solver. Problems of this kind usually require one or two major insights about the problem input to yield the problem output. The particular difficulty is usually in achieving these insights. Once they are achieved, solution of the problem becomes more or less automatic. As is true for so many distinctions, "well-defined" and "ill-defined" problem spaces are better conceived of as representing directions of a continuum rather than as representing a crisply conceived dichotomy. Solution of the former kind of problem will almost always be facilitated by one or more insights about the problem or about certain steps of the problem; solution of the latter kind of problem usually require some small steps as well as the large ones.

Problems with Well-Defined Problem Spaces

The Scope of Problems with Well-Defined Problem Spaces

In problems of this kind, it is possible to specify in some detail a problem space the traversal of which will result in a correct solution. A number of different kinds of problems with well-defined problem spaces have been studied, among them:

1. Missionaries and Cannibals Problem. The Missionaries and Cannibals Problem is one of a number of "river-crossing problems" in which a group of travelers must be transported across a river from one bank to another. What makes the task problematical is that the boat can hold only a limited number

of travelers, and that certain combinations of travelers are not permitted: In the Missionaries and Cannibals version, for example, the number of cannibals cannot be allowed to exceed the number of missionaries, since when the cannibals outnumber the missionaries, the cannibals eat the missionaries. An essentially identical problem has been studied using "hobbits" and "orcs," and very similar problems have been studied using men and elves, and silver and gold talismans. A closely related problem uses jealous husbands and wives. Problems of this general kind have been studied by Ernst and Newell (1969); Greeno (1974); Jeffries, Polson, Razran, and Atwood (1977); Reed and Abramson (1976); Reed, Ernst, and Banerji (1974); Simon and Reed (1976); and Thomas (1974). A selective review of this literature appears below.

2. Water Jugs Problem. The Water Jugs Problem can take various forms, but in all forms, the goal of problem solving is to transfer water between or among a set of jugs so as to accomplish some desired goal state. For example, one might be given a five-gallon jug and an eight-gallon jug, and be asked how it is possible to put precisely two gallons in the five-gallon jug. Either jug can be filled from a nearby sink, and water can be poured from one jug to another, but the jugs do not have gradations of measurement marked on them, and no measuring devices are available except the jugs themselves (Ernst & Newell, 1969). In a slightly different version, a mother sends her boy to the river to bring back exactly three pints of water. She gives the boy a seven-pint can and a four-pint can. The subject's task is to show how the boy can measure exactly three pints of water (Terman & Merrill, 1937). In a more difficult version of the same kind of problem, a subject is told to consider three jugs of varying capacity, e.g., jugs

A, B, and C, and is told that A will hold eight units, B will hold five units, and C will hold three units. Initially, A is full and B and C are empty. The subject's task is to determine how it is possible to divide the contents of the largest jug equally between the largest and middle-sized jugs (see Atwood & Polson, 1976; Luchins, 1942). Problems of this general kind have been studied by Atwood and Polson (1976), Ernst and Newell (1969), Luchins (1942), and Mortensen (1973).

3. Tower of Hanoi Problem. In the Tower of Hanoi Problem, the subject is presented with three pegs arranged in a linear order and *n* disks (with *n* usually about 4 or 5). The *n* disks are graded in size, and are initially stacked on the left peg, with successively larger disks closer to the bottom of the peg. The subject's task is to transfer the disks from the left-hand peg to the right-hand one. The basic constraints are that disks can be transferred one at a time from any peg (left, middle, right) to any other, that one can remove a disk only from the top position on a given peg, and that one can never place a larger disk on top of a smaller one. Various isomorphs of the basic problem have also been studied, using such vehicles as a tea ceremony and three five-handed monsters holding three crystal globes. In every case, the goal is to transfer objects in a minimum number of moves. The Tower of Hanoi Problem and its isomorphs have been studied by Egan and Greeno (1974), Ernst and Newell (1969), Hayes and Simon (1974, 1976a, 1976b), and Simon (1975).

The above list is obviously far from complete, and deals with only a limited class of problems with well-defined problem state spaces. Nevertheless, it constitutes a reasonable sampling of the kinds of problems with well-defined state spaces that have been studied.

<u>A Case Study of Problem Solving in a Well-Defined Problem Space: The Missionaries and Cannibals Problem</u>

1. <u>Nature of the problem</u>. In a typical version of the Missionaries and Cannibals problem, the subject must figure out how to transport three missionaries and three cannibals across a river. A boat is available for transportation, but it will hold only two individuals at a time. It is also possible to use the boat to transport just a single individual at a time. The number of cannibals on either side of the river can never be allowed to exceed the number of missionaries, since, in this event, the cannibals will eat the missionaries. A somewhat more difficult version of the problem uses five missionaries and five cannibals, plus a boat that will hold up to three persons at a time. Several variants of the problem have been used, all of which pose essentially the same problem to the subjects: hobbits and orcs, elves and men, and silver and gold talismans.

Performance on the Missionaries and Cannibals problem has not yet been shown to satisfy all of the criteria proposed earlier, although the means for doing so are readily available. Let us consider each criterion for task selection in turn.

Performance on the task is certainly quantifiable, and in a number of different ways. One overall measure of performance is simply the total number of moves needed to solve the problem; another overall measure is total amount of time spent in solving the problem. Each of these overall measures can be broken down further. Total number of moves can be broken down into numbers of legal and illegal moves; and number of legal moves can be further broken down into number of correct moves (those that move the subject closer

to solution) and number of incorrect moves (those that do not move the subject closer to solution). For each of these numbers of moves, one can also measure total amount of time spent on moves of that kind. Another way of quantifying performance is in terms of the number of times a given state is entered, where a state is defined by the number of missionaries and cannibals on each of the two sides of the river and by the position of the boat with respect to these two sides. Similarly, one can measure total amount of time spent in each state. These states can be subdivided, of course, in terms of whether they are legal or illegal (e.g., more cannibals than missionaries on one side of the river), and measurements can be done separately for legal and illegal states. More refined measures are also possible for particular kinds of analyses, but it should be clear at this point that quantification of performance on these problems can be done in several different ways.

No one has explicitly tested the reliability of performance in the Missionaries and Cannibals Problem. Indirect tests, however, have indicated high alternate-forms reliabilities across isomorphs (Jeffries et al., 1977) and instructional conditions (Simon & Reed, 1976).

Construct validity of performance on the Missionaries and Cannibals Problem can be inferred from a number of different sources. These sources all lead to the conclusion that performance on this task can and should be accounted for by a general theory of problem solving. Ernst and Newell (1969) showed that their General Problem Solver (GPS) program, which they took to be a theory of problem solving, was capable of solving the Missionaries and Cannibals Problem using the same basic strategies as were used in the solution

of other kinds of well-structured problems.

The construct validity of performance on the Missionaries and Cannibals Problem is also supported by the analyses of Jeffries et al. (1977), who have argued that the strategy subjects use in solving the Missionaries and Cannibals Problem is a special case of a more general strategy that can be used in solving other problems of this general kind (sometimes called MOVE problems), such as the water-jugs problem.

The one criterion for task selection on which evidence is conspicuously missing is that of empirical validity. We don't know, at this time, whether any of the indices of performance on the Missionaries and Cannibals Problem are correlated with indices of performance on other tasks of interest that are not trivially different. For example, it would be of interest to know whether Missionaries and Cannibals performance is related to IQ or to planning ability of some sort (e.g., construction of flow charts in computer programming). As has been shown earlier, empirical validity cannot be taken for granted: Tasks that one would expect to be empirically valid (e.g., animal-name analogies) sometimes show disappointing correlations with external measures.

2. Performance components. Three complete models of performance on the Missionaries and Cannibals Problem have been proposed: the GPS model of Ernst and Newell (1969), a modification of the GPS model that seems better able to account for strategies used by human subjects (Simon & Reed, 1976), and the model of Jeffries et al. (1977), which also shares certain features with GPS, but is less closely derived from GPS than is the model of Simon and Reed. I shall describe in detail here only the Simon-Reed model.

According to Simon and Reed's (1976) model, subjects may use either or

both of two basic strategies in the solution of the Missionaries and Cannibals Problem. If they use both strategies, then they do so sequentially, with strategy change occurring either suddenly (strategy shift) or gradually (strategy learning). A given strategy may incorporate both systematic and random elements. In other words, a subject may choose among alternative courses of action on the basis of a preference ordering determined by the strategy, or the subject may select one of the alternatives at random (with equal probabilities assigned to all alternatives considered). So-called "random behavior" may be interpreted as its name implies--as genuinely random behavior--or as behavior based on a mixture of other strategies not incorporated into the proposed model. The authors opt for the second interpretation, although the preferred interpretation does not affect the outcome of applying the model to data. Subjects are also assumed to seek to avoid, to a specifiable extent, reversing a move they have just made, i.e., going backward so that the problem state is what it was prior to the move that led to the current state.

The first strategy is a <u>balance</u> strategy. In this strategy, subjects select that legal move which balances the number of missionaries with the number of cannibals on each side of the river. The authors suggest that use of such a strategy is motivated by subjects' awareness that the number of cannibals cannot be permitted to exceed the number of missionaries on a given side. Subjects soon also realize as an implication of this rule that the number of missionaries cannot exceed the number of cannibals on either side, unless the opposing side has no missionaries at all (to be eaten by the cannibals).

The second strategy is a means-ends strategy. Use of this strategy entails a subject's preferring that move which takes the maximum number of persons across the river on odd-numbered moves, or the minimum number of persons back across the river on even-numbered moves. This strategy seeks to reduce as much as possible the difference between the goal state (all persons across the river) and the current state of problem solving.

3. Representation of information. Investigators studying the Missionaries and Cannibals Problem have all represented information needed and used during problem solving in terms of a problem state space. This particular state space is for a problem with three missionaries, three cannibals, and a boat that can hold a maximum of two persons. Slightly more complicated state spaces are needed for more difficult versions of the problem (e.g., five missionaries, five cannibals, and a boat holding a maximum of three persons). In Thomas's (1974) notation, each state is specified by a three-digit code, where the first digit represents the number of missionaries on the starting side, the second digit represents the number of cannibals on the starting side, and the third digit represents the location of the boat (1 if it is on the starting side, 0 if it is on the opposite side). One interesting and surprising feature of the space is its near "linearity:" At all but two states, only two legal moves exist, the correct move and a move that will result in the subject's retreating to the previous state. A branching exists at the other two states, but either branch can lead to the correct next state. The state space becomes somewhat more complicated if illegal states are added (see, e.g., Jeffries et al., p. 414), but such states are cul de sacs from

which subjects have no choice but to retreat immediately (lest they fail to solve the problem). The state space becomes more complicated in versions of the problem using more missionaries and cannibals (see, e.g., Simon & Reed, p. 88).

All investigators studying the Missionaries and Cannibals Problem have made it clear that representations of the sort they have used are "formal" representations that may or may not correspond to what subjects have in their heads. Certainly subjects are not aware of this form of representation, since they are almost never aware of the linearity of the space, and since it would not be possible for them to hold the entire space in their working memories.

Thomas (1974) was interested in testing whether subjects' actual representations corresponded to the formal representation shown above. He noted that GPS, Ernst and Newell's (1969) theory of problem solving, did in fact use the formal state space as the basis for problem-solving performance. Two separate tests of the hypothesis were made. The results indicated that the formal state space is insufficiently rich as a representation of subjects' knowledge. The information available to a subject in a given state exceeds the three items of information characterizing that state (i.e., number of missionaries and cannibals on the original side plus position of boat). Whatever the states of the problem space may represent, they are not representative of separate stages of information processing. Subjects have some kind of higher-order representation that integrates or cross-cuts stages.

Greeno (1974) reached the same conclusion as Thomas, although for a different reason. Two of Greeno's experimential conditions were a "correction condition," in which subjects were given corrective feedback if they made an error, and a "noncorrection prevent backward" condition, in which subjects were not given corrective feedback when they made errors, but in which they were

immediately informed if they made a move that took them backward in the state space. Greeno found that although performance in these two conditions was identical in the two hardest states (321 and 110), performance was worse in the former condition than in the latter in the two relatively easy states that follow the two hardest states. Greeno interpreted this result as indicating that subjects in the noncorrection condition do some looking ahead from the difficult states, but that subjects in the correction condition are not able to do this because of disruption from the corrective feedback. Thus, subjects in the noncorrection condition, at least, seemed to be organizing their responses at a level beyond that of individual states of the space.

4. Combination rules. Consider first how the basic components of problem solving in the Simon-Reed model are combined. In this model, it is hypothesized that all subjects begin solution using the balance strategy plus a "random element," and at some point switch to the means-ends strategy plus a random element. There is a certain probability of switching strategy at each move through the state space. At each move there is also a probability that a given subject will guard against returning to a previous state (which Simon & Reed refer to as an anti-loop strategy), and this probability increases over time. The probability that a subject will select his or her move according to the selected strategy rather than according to the "random element" also increases over time. Differences in problem-solving behavior as a function of experimental condition (e.g., prior practice with the Missionaries and Cannibals Problem or receipt of a hint as to how the problem should be solved) are produced by effects of the parameters of the model, namely, change in strategy-switching probability, initial probabilities of moving according to strategy rather than at random, probability of testing for a loop back to a

previous state, and the rates at which these latter two probabilities change. There are thus five parameters to be combined in the model.

Simon and Reed tested this model in two experiments. In one experiment, subjects were either given a subgoal (information regarding the placement of the boat and the numbers of missionaries and cannibals across the river) or no subgoal; in a second experiment, performance was measured on a first trial of performance and on a second trial (replication) of performance. Predictions of the model were determined through a computer simulation. The proposed model accounted for 90%, 88%, 74%, and 79% of the variance in the legal-move data in the four respective experimental conditions. The authors interpreted these data as providing a reasonable level of support for the model.

In the final paragraphs of their article, Jeffries et al. compare their model to that of Simon and Reed, and this comparison is obviously of interest here. The two models have in common their claims that people do not plan multistep sequences, that people use means-ends analysis and memory for states previously entered, and that when all else fails, people choose a move at random. The most striking difference between the models is that the Simon-Reed model assumes that subjects change strategies at some point during their experience with the problem, whereas the Jeffries et al. model assumes that a single, more complex strategy can account for problem solving throughout the course of a subject's experience with the problem. The Jeffries et al. model seems to say more about and place greater demands on memory for the problem-solving process, and it also assumes that subjects engage in later steps (i.e., stages of problem solving) only if earlier steps fail to yield a next move. Jeffries et al. note that both models account quite well for obtained

data, but that neither their data set nor that of Simon and Reed provides sufficient data to distinguish between models. Since both models seem to be accounting for the same kinds of data, it is not clear to me why this is the case. In terms of values of R^2, or percentage of variance accounted for in the data, the Jeffries et al. model does better on Jeffries et al.'s data set than Simon and Reed's model does on Simon and Reed's data set. Comparison of R^2 across experiments, however, and often even within experiments, is fraught with difficulties, and the greater predictive efficacy of the Jeffries et al. model might be due in part to its seemingly greater complication. At present, therefore, there seems to be no basis for distinguishing between the two models.

5. Durations, difficulties, and probabilities of component execution. No one has attempted to account for latencies of problem solving during solution of the Missionaries and Cannibals Problem, and hence no latency parameters have been estimated or even proposed. The parameters that have been estimated address aspects of problem solving that make various moves more or less difficult, and that affect probabilities of entering various states.

In the Simon-Reed model fitting, there were "no known systematic procedures for finding best estimates of the model's parameters in order to fit it to a set of data" (Simon & Reed, 1976, p. 90). Hence, Simon and Reed "tuned" the parameter estimates with the aid of data from the control condition of their first experiment, where subjects were asked simply to solve the Missionaries and Cannibals Problem without any hints or prior experience with the problem. "Tuning" consisted of adjusting parameters until the data from human subjects and from the simulation were almost perfectly congruent. Parameter estimates were psychologically plausible.

6. Metacomponents. Our understanding of the metacomponents of problem solving in the solution of the Missionaries and Cannibals Problem is rather minimal, and there are some metacomponents about which we know nothing.

i. Selection of performance components. The selection of performance components is motivated by several considerations on the part of the subjects. The first is the subjects' desire to attain balance between the number of missionaries and cannibals on a given side. If the numbers are not monitored, one runs the risk of creating a situation where the cannibals can eat the missionaries. In the Simon-Reed model, it is this consideration that leads to the use of a "balance" strategy early during problem solving. A second consideration is the subjects' desire to attain the final state as quickly as possible. Subjects presumably have the (correct) intuition that the Missionaries and Cannibals Problem is one that can lead to infinite looping whereby a solution is never attained, and that one way of counteracting this possibility is to keep pursuing moves that lead to closer approximations to the end state. In the Simon-Reed model, this consideration leads to the use of a "means-ends" strategy later during problem solving.

ii. Selection of representation(s). We know that subjects do not represent information about the Missionaries and Cannibals Problem solely in terms of the formal state space (Greeno, 1974; Thomas, 1974). But we do not know how subjects do represent information, nor even into what kinds of units the representation is parsed. There are some pragmatic considerations that, from the subject's point of view, would seem to place constraints upon the kind of representation that might be used. First, subjects could never hold the entire formal state-space or any analogue of it in working memory. The representation that subjects use must somehow chunk information in a way that

permits subjects to retain a local context and some sense of where they are in the global scheme of things, but that does not require retention of large numbers of states in working memory. Second, the representation must be one that is easily retrievable and modifiable. The Missionaries and Cannibals Problem requires frequent access of processes to the representation(s) upon which these processes act, and since subjects almost certainly do not have the full psychological state space represented when they start problem solving, they must be able to add to and delete from their representation on a fairly regular basis as they glean new information about the problem. Third, the representation must be one that somehow permits unitization of several pieces of information, some of which are different in kind. Obviously, the subject must be able somehow to unitize information about the number of missionaries and cannibals on a given side, and about the position of the boat. But the subject will also need some integration of this information with his or her memory of the previous state, lest the subject reenter the previous state; also, the subject must be able to hold in working memory the present and at least partial information about the previous state at the same time that the subject performs tests on the legality of the proposed state to be entered; finally, the subject must be able to remember which of the next possible states have already been tested for legality, lest the same state be tested again and again.

iii. Selection of a strategy for combining components. The balance strategy and the means-ends strategy can at best be viewed as substrategies or components embedded in the context of an overall strategy for the solution of the Missionaries and Cannibals Problem. This overall strategy includes

different elements, depending upon the model to which the subject adheres (Simon & Reed's, Jeffries et al.'s, or some other). One question that inevitably arises is that of how a subject is able to put together such a complex package of information in the absence of prior experience (for most subjects) with problems of this kind. The data of Greeno and of Thomas suggest perhaps three or four stages of information processing in solving the Missionaries and Cannibals Problem, and the models we have considered actually postulate numbers of stages at this level. It seems plausible that subjects consciously plan only three or four aspects of their information processing. The other aspects of processing that are necessary for solution of the problem may be immediate concomitants or consequents of these three or four basic aspects of planning combined with the structure of the problem. In other words, the basic decisions needed to solve the problem plus the inherent nature of the problem guide the subject into making a fairly large number of decisions that the subject may not even be aware of, or of the need for. Were this not the case, the complexity of the models, and particularly of the Jeffries et al. model, would be difficult to accept in a "performance model" of information processing. What might be the basic decisions that, once made, could lead almost automatically to the need for the remaining decisions?

First, the subject needs to decide upon an implicit "evaluation function," which in turn leads to selection of a strategy and a way of implementing that strategy. Second, the subject must decide what information is needed to start moving forward and to keep moving forward, namely, knowledge of numbers and positions of missionaries and cannibals, and of the position of the boat, for the present and the immediately preceding move. Third, the subject must decide

to check for illegality, and he or she must work out a system for doing so; this decision leads in turn to a means for selecting a next move from among the available alternatives. These basic decisions seem to force the need for all of the other decisions that will have to be made during the course of problem solving.

iv. Decision as to whether to maintain a strategy. It is obviously a matter of theoretical debate (between Simon & Reed on the one hand and Jeffries et al. on the other) as to whether subjects decide to change strategy midway through their solving of the Missionaries and Cannibals Problem. What is a clearcut decision in the Simon-Reed model, however, is a fuzzy one in the Jeffries et al. model, because of the continuous nature of the evaluation function in the latter model.

v. Selection of a speed-accuracy tradeoff. There is no evidence at all regarding speed-accuracy tradeoff. Reed (Reed & Abramson, 1976; Reed, Ernst, & Banerji, 1974) has collected latency data, but not speed-accuracy tradeoff data.

vi. Solution monitoring. It is difficult for subjects to monitor their solution processes in the Missionaries and Cannibals Problem, because they do not have the total state-space available to them, and because later states do not always resemble the final state more closely than do earlier states. At best, subjects can infer whether the drift of the states they are entering is toward the goal state, even if individual states do not always appear to be in this direction. The two most difficult states--110 and 321--appear to be ones in which extensive solution monitoring occur (Greeno, 1974; Thomas, 1974). These are the states with the maximum number of possible alternative responses (5). In the case of state 321, it is the single state in the problem (for three

missionaries and three cannibals) where it is possible to make a backward move that does not return one to the state just left. Hence, subjects seem more likely to assess at these points whether they are indeed progressing toward their goal. Thomas found that when he informed subjects at state 110 that they were "on the right track," and that the problem was "solvable from here," it improved their performance considerably. In effect, the experimenter performed the solution monitoring for the subject.

7. Problem format, problem content, and practice. Consider first the effects of problem format. The standard format for presentation of the Missionaries and Cannibals Problem has been to present subjects with the basic information required for solving the problem, and then to ask them to trace through the steps that are needed to go from the initial state to the goal state. There have been several basic variations on this format.

Thomas (1974) provided one group of subjects with feedback at state 110 that told the subjects that they were on the right track. This feedback increased the proportion of correct moves out of this state from .49 in a control group without feedback to .64 in the experimental (feedback) group, and decreased the number of backward and restarting moves from .26 to .15. Although these effects were in the predicted directions, they were not statistically significant.

Greeno (1974) had three different feedback conditions. In a first group, subjects were informed after errors that allowed orcs (the analogues to cannibals) to eat hobbits (the analogues to missionaries). In a second group, subjects were informed after errors that allowed orcs to eat hobbits, and also after moves that would produce backtracking through the state space. A third

group consisted of subjects who were also informed after both kinds of errors, but who were further informed of which response was correct. These subjects differed from the other subjects, then, in not having to experiment with other moves. Greeno found that the mean number of errors allowing orcs to eat hobbits was 9.6 in the first group, 9.9 in the second group, and 6.6 in the third group. The difference among groups was not significant, and the reduction in "eating" errors in the third group was attributed by Greeno to the subjects' being told the correct move and thereby being prevented from making more than one error in a given trial. The total number of backward moves was 12.7 in the first group, 6.7 in the second group, and 4.7 in the third group. The value for the first group differed significantly from the values for the other two groups, but the values for the other two groups did not differ significantly from each other.

Reed and Abramson (1976) performed two experiments that varied information about problem states across groups of subjects. In their first experiment, they used three missionaries and three cannibals for the test problem. Subjects received either no subgoal information, subgoal information about a subgoal that would be reached early during problem solution, or subgoal information about a subgoal that would be reached late during problem solution. The numbers of legal and illegal moves did not vary across groups, nor did solution time. Subjects given the earlier subgoal did reach that subgoal in significantly fewer moves and with significantly shorter latency than did subjects given the later subgoal; but overall performance on the problem was unaffected. Thus, the differential effect of subgoal location was limited to performance before that early subgoal was reached, and the difference was washed out when indices of performance for

the total problem were considered. In their second experiment, using five missionaries and five cannibals, subjects were given an early subgoal or no subgoal at all. The subgoal significantly reduced the mean number of legal moves from 27.6 to 20.3, significantly reduced the number of illegal moves from 5.5 to 3.7, and significantly reduced mean solution time from 883 seconds to just 437 seconds. The mean number of legal moves to the subgoal state and the time to reach that state were also significantly reduced. The authors suggested that the subgoal facilitated overall performance in the second experiment but not in the first experiment because it caused a greater reduction in the size of the state space for the second problem. Although the minimum number of moves needed to achieve a solution is the same in both versions, the number of "false" moves is far greater in the larger version of the problem, making it a more difficult problem. Thus, providing subgoal information in the larger version of the problem provides more information about moves the subject should not make. The authors conclude that a subgoal is probably not very effective in a problem space consisting of many states in which there is only one legal forward move that can be made.

As mentioned earlier, several types of problem content have been used in studies of the Missionaries and Cannibals Problem. There have been two basic kinds of manipulations. The first concerns the kinds of individuals to be transported--missionaries and cannibals, hobbits and orcs, elves and men, or silver and gold talismen. Jealous husbands and wives have been used in a problem that is similar (homomorphic) but not identical (isomorphic) to the Missionaries and Cannibals Problem, and so this variant of the problem will not be considered here. The second kind of manipulation concerns the numbers of individuals to be transported, which has been either three of each kind or five of each kind.

The data of Jeffries et al. (1977) directly address the relative difficulties of the various content isomorphs. These authors used hobbits and orcs, two variants of elves and men, and silver and gold talismen. They found no significant differences in numbers of legal moves across isomorphs. They did find significant differences in numbers of illegal moves and in numbers of errors, however. In particular, the numbers of illegal moves and errors were lowest in the hobbits-orcs condition, and highest in the silver-gold talisman condition. The two variants of the elves-men problem showed almost identical patterns of data for illegal moves and errors. The authors were able to localize the differences to the two problem states with the highest numbers of illegal moves.

The data of Reed and Abramson (1976) permit a direct comparison between the difficulty of the Missionaries and Cannibals Problem for three versus five individuals of each kind. For their control groups (standard problem presentation format), the mean number of legal moves was 20.0 in the "3MC" (three missionaries and three cannibals) group and 27.6 in the "5MC" (five missionaries and five cannibals) group; the comparable means for illegal moves were 4.1 and 5.5; the comparable means for solution time were 361 seconds and 883 seconds. Clearly, the 5MC condition was considerably more difficult than the 3MC condition.

Quite a bit of research has been done on the effects of practice upon efficacy of problem solving in the Missionaries and Cannibals Problem. Thomas (1974) was interested in part-whole transfer in problem solving. In one group, subjects simply solved the Missionaries and Cannibals Problem as presented in standard format. In a second group, subjects first solved the problem from the halfway point until the end; they then resolved the problem, except that this

second time they started the problem at its beginning. Thomas made two unexpected findings. First, initial practice on the latter half of the problem did not facilitate later performance on that part of the problem (when the second group resolved the problem), but this initial practice did facilitate later performance on the first part of the problem, that part on which the subject had not received prior practice. Second, the control group (the one that received the problem in the standard format) showed negative transfer with respect to the "part-whole" group from the first part of the problem to the second, i.e., they required more moves to solve the second part of the problem (15.5) than did the subjects who solved the second part of the problem without yet having solved the first part (12.0). Thomas's explanations of these findings were in terms of "context effects," state-specific effects, and a psychological state space that did not correspond to the formal state space. But the explanations proposed by Thomas did not go a long way toward removing the mystery surrounding these two surprising findings.

Greeno (1974) had subjects solve the hobbits-orcs problem repeatedly until subjects made no errors on two successive trials. Groups differed in feedback they received for their performance (as described earlier). Greeno found that subjects learned from positive information indicating which response was correct, rather than from elimination of errors in performance or from sampling of new strategies after commission of errors. Analysis of acquisition data was consistent with an hypothesis of all-or-none learning at individual states in the problem space, except for one state. Greeno also used an elves and men version of the problem as well as the hobbits and orcs version, and obtained essentially parallel results.

Reed, Ernst, and Banerji (1974) investigated effects of practice, although their particular focus was upon transfer between the Missionaries and Cannibals Problem and the Jealous Husbands and Wives problem. The formal state space for the problem is the same as that for the Missionaries and Cannibals Problem, if husbands are substituted for missionaries, and wives for cannibals. There is a critical difference between problems, however (which incidentally points out how the formal state space cannot capture all aspects of a problem needed in actual problem solving). In the Jealous Husbands and Wives Problem, husbands and wives are paired, such that they must always be with their own husbands, if they are with any men at all. In the Missionaries and Cannibals Problem, there is no such pairing. Reed et al. refer to the two problems as "homomorphic," meaning that there is a many-to-one mapping from the Missionaries and Cannibals Problem (where any pairing is possible) to the Jealous Husbands and Wives Problem (where only one pairing is possible). The authors sought to discover whether there would be significant transfer between problems, i.e., whether practice with one would facilitate performance with the other.

The authors conducted three experiments. In the first experiment, subjects were required to solve both problems, with half of the subjects starting with the Missionaries and Cannibals Problem and the other half starting with the Jealous Husbands and Wives Problem. In the second experiment, subjects solved the same problem twice; thus, this experiment investigated transfer within rather than across problems. In the third experiment, the procedure was identical to the first, except that the authors inserted an additional paragraph into the instructions that informed subjects of just how the two problems were related. In the first experiment, there was virtually no transfer from solution of the first problem to solution of the second problem. The second experiment was an

attempt to find out why the results of the first experiment were so disappointing. Reed et al. figured that for transfer to occur across problem types, it would have at least to occur within problem type. They therefore set out to find out if such transfer within problem type occurred. Averaged across problems, the authors found that there was a significant decrease in solution latency and in number of illegal moves, but not in total number of moves. Follow-up tests revealed that the effect was highly significant for practice on the Jealous Husbands and Wives Problem, but was only marginally significant for the Missionaries and Cannibals Problem. The second experiment showed that at least some within-problem transfer took place, and so did not isolate the reason for the failure of transfer to occur in the first experiment. The third experiment provided a way of testing whether the reason for the failure was subjects' inability to see how the two problems are related. In this experiment, the authors found that with solution latency and number of illegal moves as the dependent variables, there was significant and substantial transfer from the Jealous Husbands and Wives Problem to the Missionaries and Cannibals Problem, but not vice versa. With total number of moves as the dependent variable, there was no evidence of transfer. This experiment thus suggested that for transfer to occur, it was necessary for the more difficult problem to be presented first, and for subjects to be informed of the relationship between this problem and the less difficult one.

In an attempt to find out how transfer occurred, Reed et al. asked subjects to indicate which of four strategies best described the relationship between their strategies in solving the first and the second problem. Most subjects indicated that they "occasionally" used their memory for the first problem as a basis for solving the second problem, but that they usually attempted to solve

the second problem independently of the first. Some subjects indicated that they did not use their memory for the first problem at all, but rather solved the second problem independently of the first. Only a handful of subjects remembered "most" of their earlier moves and none remembered all.

Simon and Reed (1976), of course, were very interested in practice effects on performance, and had one set of conditions where subjects solved the Missionaries and Cannibals Problem twice in succession. They found a substantial decrease in number of legal moves from the first trial of solution to the second.

8. Individual differences within age level (adults). None of the investigators who have studied the Missionaries and Cannibals Problem have been particularly concerned with individual differences. Nevertheless, the data of Reed et al. (1974) are strongly suggestive of the existence of individual differences, at least when more than one trial is given: Subjects indicated several different levels of use of the first problem in solving the second problem. As always, individual differences may be responsible for differences in findings across the various studies that have been done.

9. Differences across age levels. The Missionaries and Cannibals Problem has not been studied developmentally, so there is no information available on developmental differences. The problem does seem susceptible to developmental investigation, however, perhaps from the secondary-school age level, upward.

10. Relationships between solution of the Missionaries and Cannibals Problem and solution of other kinds of problems. The apparent lack of interest in individual differences on the part of investigators who have studied the Missionaries and Cannibals Problem has led to a virtual absence of data regarding relationships across subjects in their ability to solve the Missionaries and Cannibals Problem

in comparison to other kinds of problems. Such individual-differences analyses would be motivated by at least two theories--those of Ernst and Newell (1969) and of Jeffries et al. (1977)-- that claim that the processes used in solving the Missionaries and Cannibals Problem are highly overlapping with the processes used to solve other kinds of problems.

Jeffries et al., for example, claim that people working on transformation (MOVE) problems such as the Missionaries and Cannibals Problem or the Water Jugs Problem

> consider only single-step move sequences, using two criteria for selecting successors: (i) select moves that lead to "better" states, where better is defined in terms of a means-ends evaluation, and (ii) avoid moves that lead to states recognized as previously visited. The details of how states are evaluated and the order in which moves are considered are specific to a particular task. (p. 436)

In the Jeffries et al. model, the memory processes are identical to those proposed by Atwood and Polson (1976) in their model of performance in the Water Jugs Problem, and the move selection process (stage model) is also very similar to that of Atwood and Polson. To the extent that there are differences between models, they are in the specifics (as opposed to the form) of the evaluation function. Such specifics would necessarily be different, since different items of information necessarily require different specific means of evaluation. Jeffries et al. compared parameters estimated from performance in their Missionaries and Cannibals isomorphs to performance obtained on the Water Jugs Problem (Atwood & Polson, 1976). Values of parameters were quite close. There is thus least tentative evidence of generality of processes across two members of the class of MOVE problems, namely, the Missionaries and Cannibals

Problem and the Water Jugs Problem. The work of Ernst and Newell (1969) suggests that the generality in methods of problem solving might extend even further. At least some correlational investigation ought to be done to determine whether patterns of individual differences, as well as parameter values, are similar across the various kinds of MOVE problems. An obvious next step in a program of research investigating the Missionaries and Cannibals Problem would be to study individual differences, and to relate them across this task, the Water Jugs task, and, perhaps, the Tower of Hanoi task.

11. Relationship between performance on the Missionaries and Cannibals Problem and intelligence. Perhaps because of the lack of interest in individual differences on the part of investigators studying the Missionaries and Cannibals Problem, no one has attempted to correlate scores for various aspects of performance on the problem with scores on any kind of general intelligence test. I believe this to be unfortunate, because there seems to be an implicit assumption in the work that performance on the Missionaries and Cannibals Problem taps at least some fundamental aspects of problem solving, and presumably, such fundamental aspects of problem solving would be important in any well-conceived notion of intelligence. An investigation of the relationship between Missionaries and Cannibals performance and measured intelligence could be an obvious part of the kind of individual-differences research mentioned above.

12. Practical relevance. The question of practical relevance has also received short shrift in the literature on the Missionaries and Cannibals Problem. I am unable to find any discussion at all in the literature

regarding what practical relevance performance on the problem might have. On the one hand, certain aspects of performance on the problem would seem to be called for in everyday problem solving: the setting of subgoals, the need to represent information in a way that moves one forward and not backward in problem solving, the use of some kind of evaluation function to choose among alternative next moves in problem solving, and so on. On the other hand, the problem seems artificial in at least some important ways: in the contrived nature of the task (regardless of which isomorph is used); in the seemingly arbitrary constraints that are placed upon accomplishment of the task (cannibals eating missionaries if they outnumber the missionaries; a boat that only holds two individuals); in the simple nature of the problem state space, in which for most moves (in the 3MC problem), there is only one legal move that can move one forward and one legal move that can move one backward; and in the clarity with which "legal" and "illegal" moves are defined. These limitations may or may not reduce or even undermine the ecological validity of the task as a representative case of real-world problem solving. Investigation of the task's ecological validity, or at least external validity of some kind, is sorely needed.

To recapitulate, the proposed metatheoretical framework can be and has been applied to the understanding of one kind of problem solving, a kind in which the problem state space is well defined. We now turn to a consideration of how the framework can be applied to the study and understanding of problem solving in an ill-defined state space.

Problems with Ill-Defined Problem Spaces

The Scope of Problems with Ill-Defined Problem Spaces.

In problems of this kind, it is difficult (at least in our present state of knowledge about problem solving) to specify in any detail a problem space the traversal of which will result in an adequate solution. Several different kinds of problems with ill-defined problem spaces have been studied, among them:

1. Hatrack Problem. In the original form of this problem (Maier, 1933), subjects are asked to construct a hatrack in an experimental room. The room (as described by Hoffman, Burke, & Maier, 1963) is eight-feet high in most places, and contains various items such as electrical conduits, lighting fixtures, fuse-boxes, beams, and minor irregularities in floor and ceiling. The only equipment explicitly made available to subjects consists of two one-inch by two-inch poles, one six feet in length and the other seven feet in length, and a three-inch C-clamp. Subjects are told that the hatrack they construct must be sturdy enough to support a heavy coat and a hat. In the more difficult version of the problem, subjects are told that they must construct the hatrack in a specified location (near the center) of the room. The solution is achieved by using the C-clamp to wedge the two poles firmly against the floor and ceiling. The two poles are allowed to overlap just enough so that they will stay firmly in place when clamped together. The clamp not only holds the poles together, but also serves as the hook on which the hat and coat can be hung. In the easier version of the problem, subjects are allowed to construct the hatrack anywhere in the room. In this case, various elements of the room can be used in fashioning a solution. Hoffman et al. (1963) classified solutions as being of five types:

Base: one board is used as a support to hold the second one vertically, with the clamp joining them at the floor level;

Balance: the boards are leaned on each other in an "X" or "T" shape with the clamp joining them;

Support: any solution using a part of the room (walls, pipes, ceiling beams, etc.) to hold up the construction (limited to the use of the ceiling pipes and beams during the test problem);

Ceiling Suspension: boards are wedged between ceiling and ceiling pipes or ceiling beam, with clamp joining them or appended to the joined ends;

Floor-Ceiling: the correct solution in which the boards are wedged between floor and ceiling and joined tightly by the clamp.

This problem (in its two versions as well as minor variants of them) has been studied by Burke and Maier (1965); Burke, Maier, and Hoffman (1966); Hoffman, Burke, and Maier (1963); Judson, Cofer, and Gelfand (1956); Maier (1933, 1945, 1970); Maier and Burke (1966); Raaheim (1974); and Saugstad (1955).

2. Two-String Problem. In the original form of this problem (Maier, 1931), subjects are brought into a large room containing many objects, such as poles, ringstands, clamps, pliers, extension cords, tables, and chairs. The experimenter hangs two cords from the ceiling. One hangs near the center of the room, the other near a wall. The cords are of sufficient length to reach the floor. Subjects are told that their task is to tie the ends of the two strings together. It soon becomes apparent to subjects that the cords

are far enough apart so that it is not possible to hold both cords in one's hands simultaneously. Subjects must therefore use the materials in the room to attain a solution to the problem. This problem and its variants have been studied by Duncker (1945), Maier (1930, 1931, 1933, 1945, 1970), Maier and Burke (1966), Maier and Janzen (1968), Raaheim (1974), and Saugstad (1955, 1957, 1958).

3. Radiation Problem. This problem, originally proposed by Duncker (1926), is usually posed in the following form:

> Given a human being with an inoperable stomach tumor, and rays which destroy organic tissue at sufficient intensity, by what procedure can one free him of the tumor by these rays and at the same time avoid destroying the healthy tissue which surrounds it? (Duncker, 1945, p. 1)

Proposals for solving this problem are usually of three basic kinds (Duncker, 1945). One kind attempts to avoid contact between the rays and the healthy tissue. For example, subjects might suggest that the rays be sent down a free path to the stomach, such as the esophagus; that healthy tissue be removed from the path of the rays, as by inserting a cannula; that a protective wall be inserted between the rays and the healthy tissue; or that the tumor somehow be displaced toward the surface, as by pressure. A second kind of solution attempts to desensitize the healthy tissue. For example, subjects might

suggest that a desensitizing chemical be injected into the tumor victim; or that the victim be immunized by adaptation to weak rays. The third kind of solution attempts to lower the intensity of the rays on their way through healthy tissue. The preferred answer is of this kind, namely, that weakened rays originating from several different sources be sent through the body such that the rays all converge upon the tumor. At this point, and only at this point, will the rays be of sufficient intensity to destroy tissue, which in this case will be tumor tissue. The radiation problem has been studied by Duncker (1926, 1945) and by Gick and Holyoak (in press).

These three examples of problems with ill-defined problem spaces provide only a minimal sampling of the problems of this kind that have been studied. They are sufficient, however, to permit a contrast to the kind of problem considered earlier, that with a well-defined problem space. There are several salient differences between the two kinds of problems. First, problems with ill-defined problem spaces seem to depend for their solution upon the attainment of a single major insight. Indeed, problems of this sort are often referred to as "insight problems." Problems with well-defined problem spaces seem to depend for their solution upon the attainment of a sequence of relatively more minor insights. No one striking realization marks the difference between success and failure in problem solving, as it can in insight problems. Second, in problems with well-defined problem spaces, it is possible to represent the problem space in terms of a sequence of discrete and well-articulated states. It is this property that leads to the problems being referred to as having "well-defined problem spaces." In problems with ill-defined problem spaces, it is not possible to represent the problem space in terms of a sequence of discrete and well-articulated states. This property is the one that leads to the problems being referred to as having

"ill-defined problem spaces." Third, in problems with well-defined problem spaces, the end-state is similar or identical in kind to the starting state. For example, in the Missionaries and Cannibals Problem, both the starting and end states specify numbers of missionaries and cannibals on each side of a river bank. The two states differ only in the number of each kind of individual on each side. In problems with ill-defined problem spaces, the end-state is different in kind from the starting state. For example, in the Hatrack Problem, the end-state posits the existence of a hatrack. The hatrack does not yet exist in the starting state, nor is it clear how the input in the starting state can be transformed to create a hatrack.

"Insight" problems were a popular subject of study for Gestalt psychologists, whose major concerns (such as the specification of the circumstances under which "insight" occurs) differed in many respects from those of modern-day information-processing psychologists. As a result, much of the research that was done on insight problems was addressed to questions that no longer seem terribly interesting today; and many of the questions that do seem interesting were simply never addressed. Since the reviews presented in this work are guided by theoretical questions purported to be of interest to modern-day psychologists, much of the discussion presented below will propose what needs to be studied, rather than reviewing what has already been studied that is not of contemporary interest.

A Case Study of Problem Solving in an Ill-Defined Problem Space: The Hatrack Problem

1. Nature of the problem. The Hatrack Problem requires experimental subjects to construct a hatrack out of two poles of unequal length and a C-clamp. In the easier version of the problem, subjects are allowed to use

various structural features in the experimental room (ceiling beams, lighting fixtures, fuse-boxes, etc.) to aid in construction of the hatrack. In the harder version of the problem, nothing can be used except the given elements. In this case, the solution is attained by connecting the poles with the C-clamp and wedging them against the floor and ceiling. The C-clamp is used as the hook on which to hang a hat and coat.

Consider how performance on the Hatrack Problem meets the various criteria proposed earlier. First, performance on the problem can be quantified in a number of different ways. These include time to solution, proportion of solutions that meet the constraints originally set out by the problem, probability distribution of various solutions, and proportion of subjects proposing any solution at all.

Second, it is unfortunately difficult or impossible to measure the reliability of most, but not all, indices of performance on the Hatrack Problem. Test-retest reliability cannot be feasibly measured, because once a person has solved the Hatrack Problem, it is spoiled as a future measure of problem solving skill. In this problem, once the solution is obtained, it is trivially easy to obtain the solution in subsequent trials on the problem. It is also unclear as to how, if at all, internal consistency reliability could be measured, since, in general, measures of performance are available only for performance in the final state. It would be possible to measure reliability for performance on insight problems in general, as opposed to one specific insight problem, by constructing a test that consisted of multiple problems of this kind, and by computing internal consistency of performance on such a test.

Third, there are data supporting a favorable assessment of the con-

struct validity of performance on the Hatrack Problem. Investigations of the problem have served several theoretical purposes:

1. The problem has been used in investigations of whether problem solving of the kind required by the Hatrack Problem can be understood solely in terms of reproductive thinking, or whether it must be understood in terms of productive thinking as well. A major advocate of the former position, Saugstad (1955), would argue that individual differences in problem solving can be understood solely in terms of past learning of the elements or functions needed for solving a given problem. Availability of these functions is sufficient for problem solving in the new situation. In effect, the problem is solved by a "mechanism of equivalent stimuli." A major advocate of the latter position, Maier (Hoffman, Burke, & Maier, 1963; Maier, 1933, 1945; Maier & Burke, 1966), would argue that individual differences in problem solving must be understood in terms of reasoning with past learning as well as in terms of the learning itself. The ability to combine previously learned elements is critical to solution of a given problem. Raaheim (1974) has taken a position intermediate between these two, although closer to Maier's. According to Raaheim, problem solving is an activity in which an individual attempts to dispense with deviating elements in a problem situation in order to make the new problem situation equivalent to situations encountered in the past.

2. Hoffman, Burke, and Maier (1963) used the Hatrack Problem to investigate whether an experimenter's (positive or negative) evaluations of subjects' performance on an earlier and easier problem affects their performance on a later and more difficut problem. The easier problem was the easier version of the Hatrack Problem and the more difficult problem problem was the harder version of the Hatrack Problem.

3. Burke, Maier, and Hoffman (1966) used the Hatrack Problem to study the question of what makes for a good hint in problem solving. Hints varied in whether they were given before or after problem solving began. A major purpose of the investigation was to discover and classify the various functions hints can serve in facilitating and impeding problem solving.

4. Maier (1933) used the Hatrack Problem to test whether instructions on overcoming ingrained sets and habits could facilitate problem solving. Experimental subjects were given instructions urging them to do things such as "keep your mind open for new combinations and do not waste time on unsuccessful attempts" and "do not be a creature of habit and stay in a rut. Keep your mind open for new meanings." Control subjects were not given instructions of this kind.

Fourth, we need to consider the empirical validity of performance on the Hatrack Problem. Evidence (to be presented later) is scanty and only modestly encouraging. But the tests of empirical validity that have been performed (Burke & Maier, 1965) have been so weak that one must be hesitant to draw any conclusions solely on the basis of the previously established results.

To conclude this section, there is at least some justification for the study of the Hatrack Problem as an index of problem-solving ability or skill. The evidence supporting the usefulness of the problem as an object of study is weaker than that for other kinds of problems (or reasoning items) we have considered. The weaker evidence must be viewed in the context of the fact that most investigations of the problem were done a number of years ago, when different theoretical questions were of primary interest from those that are

of primary interest today. I would view the evidence as incomplete rather than as unfavorable. The major questions one might today like to have answered remain unanswered and for the most part unasked.

2. Performance components. There is no research identifying components of information processing in the Hatrack Problem. I am prepared to speculate, however, that the performance components involved in solution of this problem are highly overlapping with, if not identical to, those involved in certain forms of inductive reasoning, such as reasoning by analogy. Indeed, Maier's view of problem solving in insight problems might be viewed as one of problem solving by analogy. The major difference between insight problems and standard analogy problems would seem to be that the components are much more difficult to apply to insight problems than they are to standard analogy problems. In the analogies, the structure of the problems is clearly defined, whereas in the insight problems it is not.

The subject must first encode the problem as it is posed, and the materials that are presented to the subject as means to solve the problem. These materials include the two poles and the C-clamp. Many, if not most subjects will not initially encode as relevant two critical elements of the problem solution, namely, the floor and the ceiling of the experimental room. Next, the subject must infer how elements of hatracks with which he or she has been familiar have functioned in these previously known hatracks. These elements must then be mapped onto the elements of the present situation. If no mapping is immediately available, as will most likely be the case, then it will be necessary for the subject to map elements of hatrack-like structures onto the current situation in an attempt to find elements from analogous struc-

tures (flag poles, pole lamps, etc.) that can be mapped onto the elements of the current situation. Once the mapping is completed (and this is almost certainly the most difficult operation to complete), the subject must figure out how to <u>apply</u> the current elements in a way that is analogous to that inferred for past elements so that the present elements can also be combined into a hatrack. In the event that the subject generates multiple possible solutions, the subject must <u>compare</u> them and decide which is most viable. The subject must then attempt to <u>justify</u> the best (or only) solution as close enough to an ideal to be minimally acceptable. If the solution is acceptable, the subject <u>responds</u> with it. Otherwise, the subject must try to find another solution, repeating earlier problem-solving operations.

This view of problem solving during solution of the Hatrack Problem contains within it both a theoretical implication and a practical implication. The theoretical implication is that problem solving in insight problems (i.e., problems with ill-defined problem spaces) is primarily analogical. The problems are particularly difficult to solve because the subject must perceive some very nonobvious relationships. The basic terms of the analogy are ELEMENTS OF HATRACKS AND HATRACK-LIKE STRUCTURES I HAVE KNOWN : HATRACKS AND HATRACK-LIKE STRUCTURES I HAVE KNOWN :: ELEMENTS OF THE PRESENT SITUATION : A NEW HATRACK (the nature of which has to be figured out). The practical implication is that problem solving in insight problems can be studied in ways comparable to those used for studying reasoning by analogy. The method of precueing, in particular, would seem to be relevant. Subjects could be precued with information sufficient for performing various operations (encoding, inference, mapping, application, comparison, justification) and

combinations of operations, and the effects of these precues on ease or difficulty of problem solving can then be assessed. The method could be used with precueing information providing needed knowledge for performing each of the successive operations (as has been done in reasoning by analogy), or it could be used with precueing information providing needed knowledge for performing combinations of operations that are not necessarily successive in information processing.

3. Representation of information. None of the research that has been done on the Hatrack Problem has explicitly addressed the question of how information is represented in memory. Some form of representation is needed that can account for people's ability to draw analogies to past experiences in order to figure out how the elements of the experimental situation can be combined into a hatrack. In particular, the representation must be able to account for the fact that the elements of the present situation have never been combined in this particular way before, and no previous hatrack has ever been encountered that was constructed of just these elements.

Schank (1979) has proposed a kind of memory structure that expands upon earlier ideas from organization theory (Bower, 1971; Tulving, 1966) and that seems suitable for the present purpose. In this structure, information is stored in the form of "memory organization packets," or MOPs. "The purpose of a MOP is to provide expectations that enable the prediction of future events on the basis of previously encountered structurally similar events. . . . The ability of MOPs to make useful predictions in somewhat novel situations for which there are no specific expectations but for which there are relevant experiences from which generalized information is available, is crucial to

our ability to understand" (Schank, 1979, p. 46).

A MOP might be expected to exist for the hanging of a hat on a hatrack. According to Schank, a given MOP will usually have "strands" corresponding to (a) reasons for the MOP existing if it's a state, or reasons for doing the MOP if it's an action; (b) enabling conditions for the state or action; (c) results of doing the MOP if it's an action; (d) normative methods of achieving or satisfying a given state; (e) what goals the state or action relates to and which it affects; (f) associated states; and (g) associated actions. Accessing the MOP for "hanging a hat on a hatrack" will probably not enable one to solve the Hatrack Problem. The subject presumably must enter a MOP for some associated state or action in order to access the elements necessary for creating a hatrack. For example, a likely associated state is a flag pole or pole lamp, the latter of which is supported by its tight fit to both floor and ceiling. The analogy to the present situation might provide the clue for solution of the problem.

4. <u>Combination rules.</u> Specific alternative strategies for solving the Hatrack Problem have not been explicitly investigated in previous research. An analysis of the task situation, however, suggests at least several plausible approaches to the problem, any one or combination or which might be used in solving the Hatrack Problem.

1. <u>Focusing upon elements of present situation</u>. In this strategy, the subject focuses upon the elements of the present situation, and tries to conceive of how these elements might somehow be combined to create a hatrack. The subject would thus think about what uses a C-clamp and a pair of poles might have in creating a hatrack.

2. Focusing upon elements of previously encountered hatracks. This strategy entails the subject's focusing upon elements of previously encountered hatracks. The strategy is in some respects the opposite of the first one. In the first strategy, the subject tries to relate the new elements to elements of old hatracks. In this second strategy, the subject tries to relate old elements of previously known hatracks to elements of the new hatrack. The subject may reflect upon hatracks he or she has known, trying to find one that is of a construction that might be roughly suitable in the present instance.

3. Focusing upon elements of prototypical hatracks and their variants. In this strategy, one frees oneself from specific past instantiations of hatracks, and tries to construct one or more hatracks that are prototypical, but that do not correspond to any specific hatracks one has previously seen.

4. Focusing upon a receptive state of mind. A fourth strategy is to try to clear one's mind of any particulars at all, and to attain a receptive state of mind. In this strategy, the subject essentially waits for a flash of insight to strike. Information processing, to the extent that it exists at all during this strategy, is below consciousness and not subject to introspective report.

Strategy usage and effectiveness might be inferred in at least two different ways. One way would be to have subjects think aloud as they solve the Hatrack Problem, and attempt to classify their strategy usage on the basis of the protocols thereby obtained. Effectiveness of the strategies could be inferred by noting how often each strategy leads to an acceptable

solution. A second way would be to train subjects to use particular strategies or combinations of strategies, and to compare rates of success in problem solving with those for subjects who are untrained. Presumably, training subjects to use the strategy they are already using should have no differential effect upon success rate, whereas training subjects to use alternative strategies presumably is likely to have some differential effect. This paradigm also permits a direct comparison of the effectiveness of the various strategies that have been trained.

5. Durations, difficulties, and probabilities of component execution. No one has tested any information-processing models of performance on the hatrack problem, and, to my knowledge, the sketchy model proposed earlier (in items 2-4) is the only model that has been proposed. Maier (1945; Hoffman, Burke, & Maier, 1963) has analyzed problem difficulty in more global respects, however. Maier (1945) found that 12 of 25 subjects with no prior experience on the hatrack problem or similar problems were able to solve the hatrack problem in 30 minutes. Hoffman et al. (1963) also studied subjects with no prior experience, and broke down performance by time to solution. They found that of 30 subjects given 30 minutes to solve the hatrack problem, 8 solved the problem within 5 minutes, 13 solved it within 10 minutes, 14 solved it within 15 minutes, 15 within 20 minutes, 15 within 25 minutes, and 15 within 30 minutes. Thus, almost all subjects who reached a satisfactory solution did so within 10 minutes.

6. Metacomponents. Hoffman et al. examined numbers of subjects who proposed each of five different solutions (presumably reflecting different strategies) during the first 10 minutes of experience with the problem. The

first four strategies for making a hatrack were deemed to be incorrect; the last was deemed to be correct. Since subjects could attempt more than one solution, the sum of the number of attempts exceeds 30 (the number of subjects). Of the 30 subjects, 16 proposed a "base" solution, in which one board was used as a support to hold the second one vertically, with the clamp joining the boards at the floor level; 16 proposed a "balance" solution, in which the boards were leaned on each other in an "X" or "T" shape with the clamp joining them; 7 proposed a "support" solution, in which a part of the room (e.g., walls, pipes, ceiling beams, etc.) was used to hold up the construction; 8 proposed a "ceiling suspension" solution, in which boards were wedged between ceiling and ceiling pipes or ceiling beam, with the clamp joining them or appended to the joined ends; and 13 proposed a "floor-ceiling" solution—the correct solution—in which the boards were wedged between floor and ceiling and joined tightly by the clamp.

Raaheim (1974) examined numbers of subjects who repeated unsuccessful attempts at solution various numbers of times. He found that of 37 subjects who never reached a solution (out of a total of 60 subjects), 9 subjects repeated unsuccessful attempts from 1 to 4 times, 19 subjects repeated unsuccessful attempts from 5 to 8 times, and 9 subjects repeated unsuccessful attempts from 9 to 14 times. Raaheim queried as to why presumably intelligent university students would repeat time and again solutions that had been designated by the experimenter, "unsuccessful." He concluded that "the most likely answer is that they are trying to solve a task <u>other</u> than the one intended by the experimenter. While the instructions aim at <u>some construction</u>, a nameless, unusual, but sturdy and quite ingenious sort of thing, the subjects nearly all very in-

tensively try to find a way of replacing the type of hatrack they know of from their past experience" (p. 49). Thus, these subjects fail in their initial definition or conceptualization of the problem task.

7. Problem format, problem content, and practice. Two major variations have been attempted in problem format. One variation is where in the room the hatrack must be constructed. In the easier variant, the hatrack can be constructed anywhere in the room, so that the subject can make use of various features of the room to facilitate construction. In the harder variant, the hatrack must be constructed in the center of the room, so that only the clamp and the two poles are available. A second variation is in whether the problem is presented with the actual materials or in written form. Raaheim (1974) found that the proportion of subjects writing down the correct solution (4/64) was about the same as the proportion of subjects choosing the correct solution first when using the actual materials (4/60).

Since the hatrack problem is content-bound (i.e., it is about a hatrack), no alternative contents have been explored.

Practice effects of various types have been widely studied in the literature on the hatrack problem. Research on practice effects has taken several different forms. Most of this research has dealt not with practice effects per se, but with whether "availability of functions" necessary for solution of a problem such as the hatrack problem is sufficient to guarantee solution of that problem. On the one hand, Saugstad (1955) has suggested that if a subject has available all of the functions (items of knowledge) necessary to solve a problem, then solution will be more or less automatic. On the other hand, Maier (Hoffman, Burke, & Maier, 1963) has argued that availability

of the necessary functions for solving a problem is not sufficient to guarantee solution of the problem: The subject may or may not be able to put together these functions into a workable strategy for solution.

One way of exploring this issue has been through a transfer paradigm. Hoffman et al. assigned 90 subjects at random to one of three conditions. In a no experience condition, subjects were given the difficult version of the hatrack problem (construct it in the center of the room) immediately upon entering the experimental room. In two prior experience conditions (varying in type of reinforcement), subjects were given the easy version of the problem first (construct it anywhere in the room), and were encouraged to construct as many different types of hatracks as they could. Subjects with prior experience on the easier version of the problem performed significantly worse than subjects with no prior experience: Whereas only 25% of the subjects in the former groups solved the problem, 50% of the subjects in the latter group did. Hoffman et al. interpreted these results as showing that providing subjects with a great variety of functions can inhibit problem-solving performance by establishing misleading problem-solving sets. Thus, whereas the prior-experience groups must have had at least as many functions as the group with no prior experience, they performed more poorly because the correct functions were not automatically utilized.

Maier (1945) used a somewhat different transfer paradigm. In his experiment, 75 subjects were equally divided into three groups, all of which were asked to solve the hatrack problem. In one group (a control group), subjects tried to solve the hatrack problem as soon as they entered the experimental room. In a second group, subjects were asked to help build two

structures that could be used in the two-string problem and that could each be used as a hatrack. Subjects were told that the purpose of building this structure was to get them adjusted to the real problem situation. Subjects were taken to another end of the experimental room, and were asked to construct a hatrack in a certain spot. The "two-string" structures were left standing. In the third group, procedures were the same as in the second group, except that the "two-string" structures were disassembled before the subject was asked to build the hatrack. Performance was best in the group in which subjects were shown the two-string structure and in which the structure was left standing (18/25 reached solution). Performance was intermediate in the group in which subjects were shown the two-string structure and in which the structure was disassembled (12/25 reached solution). Performance was worst in the group in which subjects were *not* shown the two-string structure (6/25 reached solution). Thus, in this experiment, prior experience helped, but was clearly not sufficient for solution of the hatrack problem. There were still substantial numbers of subjects who did not reach a solution, despite prior experience that made available to them all of the functions needed for solution of the hatrack problem.

Another paradigm for studying effects of availability of functions is one that employs hints toward problem solution. Maier and Burke (1966) used one such paradigm. Subjects (135 male college students) were initially given 15 minutes to solve the hatrack problem. Those failing to solve the problem were given one of two hints. One hint informed the subjects that the ceiling of the room was part of the solution; the other hint informed the subjects that the clamp must serve as the hat (or coat) hook. Subjects were then

given an additional 20 minutes to solve the problem. After this portion of the experiment ended, subjects were given an "availability of functions" test. They were taken into another experimental room in which the correct floor-ceiling solution to the Hatrack Problem had been constructed. Each subject was given a piece of paper on which to list as many functions or uses of the "structure" before him as he could possibly think of. Subjects were given five minutes to complete this task. Fifty subjects who had previously failed to solve the Hatrack Problem were then returned to the first experimental room, and again asked to construct a Hatrack. Of the 135 subjects, 51 solved the problem without a hint; 34 solved the problem after a hint was given; and 50 never solved the problem. Thus, the hint did seem to facilitate performance. Subjects in the three groups were strikingly similar in their "available functions." All but one of the 135 subjects recognized the structure as a potential hatrack, coatrack, hangar, etc. But of the 50 subjects who failed to solve the Hatrack Problem initially and were then asked to solve it after the availability of functions test, 7 were still unable to solve the problem, despite the fact that all 7 had listed the necessary function for the structure they had seen in the availability of functions test. Thus, for these subjects, at least, availability of functions was insufficient to guarantee solution of the hatrack problem.

8. Individual differences within age level (adults). Evidence has already been cited to the effect that individuals differ in their success in solving the Hatrack Problem, and in the solutions they propose. Maier (1933) found a significant sex difference in success in problem solution: Men performed significantly better than women. Maier (1945) replicated this difference,

and found that the difference held up without regard to whether or not subjects had prior experience with the "two-string" type of structure before solving the hatrack problem.

9. Differences across age levels. I have been unable to locate any developmental investigations of performance on the hatrack problem.

10. Relationships between peformance on the Hatrack Problem and on other problems. No direct correlational studies appear to have been carried out relating performance on the Hatrack Problem to performance on other insight problems, such as the two-string problem and the water-jugs problem. I would argue that the processes, representations, and combination rules described earlier (items 2-4) would be applicable to these and other types of insight problems. Hence, I would expect performances on the various problems to be about as highly intercorrelated as the probably not very high reliabilities of the performances would allow. Burke and Maier (1965) correlated success on the Hatrack Problem (evaluated simply as pass-fail) with success on various kinds of pencil-and-paper tests measuring skills that would loosely fall into the problem-solving domain: ideational fluency, spontaneous flexibility, adaptive flexibility, redefinition. Only one of seven correlations was significant, and this maximal correlation of .19 was scarcely impressive in magnitude. Raaheim (1974) found that level of activity in the Hatrack Problem correlated significantly with success in solving the problem, but nevertheless claimed that the Hatrack "Problem" is not a problem at all. Recall that according to Raaheim, a problem situation is "the deviant member of a series of earlier situations of the same sort" (p. 22).

> It may be argued that the Hatrack situation must not be looked upon by the subjects as a problem of how to build a more or less ordinary looking hatrack by some extraordinary means. Rather it must be looked upon as a task of constructing something quite different from what is usually used for hanging up coats. But then, if it is not the problem of making a hatrack, is the situation facing the individual any problem situation at all? Is there any series of situations from the past to which the present one may be said to belong, i.e., a series of situations that fits in with the solution wanted by the experimenter? If not, the Hatrack task does not fall within the category of tasks encompassed by our definition of problem situations. (p. 49)

Raaheim claims that the Hatrack situation can be turned into a problem situation by giving subjects one or more hints that relate the situation to previous situations with which they are familiar.

11. Relationship between problem solving on the Hatrack Problem and intelligence. Burke and Maier (1965) correlated performance on the Hatrack Problem (success vs. failure in solution) with scores on the verbal and mathematical sections of the Scholastic Aptitude Test. The correlations were trivial (each was -.04). These are the only correlational data that I have been able to find. On the one hand, they fail to support the notion that there is any relationship between Hatrack problem-solving and measured intelligence. On the other hand, the measure of performance is so crude (pass-fail) and the range of student ability probably so restricted (subjects were University of Michigan undergraduates) that Burke and Maier's test of the

relationship between problem solving and intelligence seems wholly inadequate.

12. Practical relevance. Maier (1933) investigated whether it is possible to improve performance on insight problems, including the Hatrack Problem, by giving subjects a prior lecture on problem-solving skills. The lecture covered 13 points, three of which were specific hints on how to solve the problems:

> (1) Locate a difficulty and try to overcome it. If you fail, get it completely out of your mind and seek an entirely different difficulty.
>
> (2) Do not be a creature of habit and stay in a rut. Keep your mind open for new meanings.
>
> (3) The solution-pattern appears suddenly. You cannot force it. Keep your mind open for new combinations and do not waste time on unsuccessful attempts. (p. 147).

The training was successful in significantly improving problem-solving performance. It was approximately equally beneficial for good and poor reasoners, but was more beneficial for women than for men.

The Hatrack Problem has been studied almost exclusively by psychologists whose theoretical concerns differ considerably from the ones proposed here to be of major importance, and from the ones that concern most contemporary information-processing psychologists. The practical relevance of previous research on the Hatrack Problem does not appear to be particularly great, but one cannot thereby infer that research could not be done on the problem that would have greater practical relevance. Research into the nature of insight, into the generalizability of the components, representations, and

strategies used in solving the Hatrack Problem, and into metacomponential decision-making in problems with ill-defined problem spaces would all seem to have potential practical relevance. Because most problems encountered in the real world do have ill-defined problem spaces, it seems that research into such problems may eventually have greater practical payoff than research into the more tractable problems with well-defined problem spaces.

REASONING, PROBLEM SOLVING, AND INTELLIGENCE

I have reviewed in this chapter only a small segment of the literature that could sensibly be viewed as dealing with "reasoning, problem solving, and intelligence." But I believe the literature I have reviewed is fairly representative of work in the field. If there are biases in coverage, and almost certainly there are, they are probably toward greater coverage of work that I consider to be more theoretically motivated and that is concerned more with how reasoning and problem solving relate to general aspects of cognition and intelligence. A major purpose of the chapter has been to support the view that the interface between reasoning, problem solving, and intelligence can be profitably pursued in terms of answers to twelve questions; these answers seem to constitute a reasonably coherent and complete account of the psychological phenomena of interest in a given domain of research. Obviously, these are not the only questions that might be posed, and some of these might be combined or deleted. But the set seems to work reasonably well in generating coverage of a task domain.

Intelligence is an amorphous concept, but if one accepts a global definition of it as adaptability to the varied situations in which one may find oneself, then the study of reasoning and problem solving appears to provide a good entree into the study of intelligence, because nontrivial adaptation inevitably will require reasoning and problem solving in various forms and guises. At least some of these forms and guises are proposed to draw upon the components, strategies, and representations considered in this chapter.

Not all reasoning and problem-solving tasks seem to be equally good

measures of intelligence. Raaheim (1974) has proposed that problems of intermediate difficulty appear to be the best measures. My own emphasis would be somewhat different. I would claim that the best tasks to study are those that are "nonentrenched," but that rely on processes, representations, and strategies shared with real-world tasks. By nonentrenched tasks, I mean ones that require strategy planning and execution of a kind that requires nonroutine kinds of thinking and behaving. My view, for which I claim no originality, is that intelligence is in large part the ability to acquire and think with new conceptual systems and to solve novel kinds of tasks.

This view seems consistent with many of our everyday notions about intelligence, if not with all of our research about it. A student is likely to be considered more intelligent if he or she can master a new kind of course (say, calculus or foreign language) than if he or she can master another course that differs in substance but not in kind from courses the student has taken previously. The student is likely to be considered more intelligent if he or she can solve new kinds of problems, rather than merely if he or she can solve problems very similar to those that have been encountered numerous times in the past.

On this view, a problem such as the Hatrack Problem should seem to provide an excellent way of measuring intelligence, and yet there is no evidence that this is the case. In fact, I doubt that the Hatrack Problem does provide a very good measure of intelligence in the usual sense of the term. On the one hand, I believe that the attainment of insights into novel kinds of problems is an essential ingredient of intelligence. On the other hand, I doubt that

there is any general ability that could be labeled "insight." Different people seem to have their best insights into different kinds of problems. And the class of problems represented by the constructions of a hatrack from two poles and a C-clamp is probably not a particularly interesting class of problems in terms of which to study people's insights. On this view, then, ecological validity of content is potentially of great relevance. If one wishes to study a scientist's or a business executive's insights, one would do best studying them in the domain to which they are normally applied in that person's day-to-day environment. Probably some tasks (such as analogies) measure intellectual functioning of such a basic kind that ecological validity is less important. But if one's goal is to study the ability to handle new kinds of situations successfully, then one probably should make sure the situations are both new and of the kind that a given person will be likely to encounter. Performance in such situations seems to be exactly of the kind that should be studied by those interested in the interface of reasoning, problem solving, and intelligence.

References

Ace, M. C., & Dawis, R. V. Item structure as a determinant of item difficulty in verbal analogies. Educational and Psychological Measurement, 1973, 33, 143-149.

Achenbach, T. M. Standardization of a research instrument for identifying associative responding in children. Developmental Psychology, 1970, 2, 283-291. (a)

Achenbach, T. M. The children's associative responding test: a possible alternative to group IQ tests. Journal of Educational Psychology, 1970, 61, 340-348. (b)

Achenbach, T. M. The children's associative responding test: a two-year followup. Developmental Psychology, 1971, 5, 477-483.

Ajzen, I. Intuitive theories of events and the effects of base-rate information on prediction. Journal of Personality and Social Psychology, 1977, 35, 303-314.

Anderson, J. R. Language, memory, and thought. Hillsdale, N.J.: Erlbaum, 1976.

Atwood, M. E., & Polson, P. G. A process model for water jug problems. Cognitive Psychology, 1976, 8, 191-216.

Begg, I., & Denny, J. Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning. Journal of Experimental Psychology, 1969, 81, 351-354.

Bower, G. H. Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), Adaptation-level theory. New York: Academic Press, 1971.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1). Hillsdale, N.J.: Erlbaum, 1978.

Brown, A. L., & DeLoache, J. S. Skills, plans and self-regulation. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Erlbaum, 1978.

Bryant, P. E., & Trabasso, T. Transitive inferences and memory in young children. Nature, 1971, 232, 456-458.

Burke, H. R. Raven's Progressive Matrices: A review and critical evaluation. Journal of Genetic Psychology, 1958, 93, 199-228.

Burke, R. J., & Maier, N. R. F. Attempts to predict success on an insight problem. Psychological Reports, 1965, 17, 303-310.

Burke, R. J., Maier, N. R. F., & Hoffman, L. R. Functions of hints in individual problem-solving. American Journal of Psychology, 1966, 79, 389-399.

Burt, C. The development of reasoning in school children. Journal of Experimental Pedagogy, 1919, 5, 68-77.

Carroll, J. S., & Siegler, R. S. Strategies for the use of base-rate information. Organizational Behavior and Human Performance, 1977, 19, 392-402.

Case, R. Intellectual development from birth to adolescence: A neo-Piagetian interpretation. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Erlbaum, 1978. (a)

Case, R. Piaget and beyond: Toward a developmentally based theory and technology of instruction. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1). Hillsdale, N.J.: Erlbaum, 1978. (b)

Cattell, R. B., & Cattell, A. K. S. Test of g: Culture Fair, Scale 3. Champaign, Illinois: Institute for Personality and Ability Testing, 1963.

Ceraso, J., & Provitera, A. Sources of error in syllogistic reasoning. Cognitive Psychology, 1971, 2, 400-410.

Chapman, L. J. Illusory correlation in observational report. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 151-155.

Chapman, L. J., & Chapman, J. P. Atmosphere effect re-examined. Journal of Experimental Psychology, 1959, 58, 220-226.

Chapman, L. J., & Chapman, J. P. Genesis of popular but erroneous psychodiagnostic

observation. Journal of Abnormal Psychology, 1967, 72, 193-204.

Chapman, L. J., & Chapman, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 1969, 74, 271-280.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press, 1965.

Clark, H. H. The influence of language in solving three-term series problems. Journal of Experimental Psychology, 1969, 82, 205-215. (a)

Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404. (b)

Clark, H. H. More about "Adjectives, comparatives, and syllogisms": A reply to Huttenlocher and Higgins. Psychological Review, 1971, 78, 505-514.

Clark, H. H. Difficulties people have answering the question "Where is it?" Journal of Verbal Learning and Verbal Behavior, 1972, 11, 265-277. (a)

Clark, H. H. On the evidence concerning J. Huttenlocher and E. T. Higgins' theory of reasoning: A second reply. Psychological Review, 1972, 79, 428-432. (b)

Clark, H. H. Semantics and comprehension. In T. A. Sebeok (Ed.), Current trends in linguistics (Vol. 12): Linguistics and adjacent arts and sciences. The Hague: Mouton, 1973.

Copi, I. M. Introduction to logic (fifth edition). New York: Macmillan, 1978.

Dawis, R. V., & Siojo, L. T. Analogical reasoning: A review of the literature. Effects of social class differences on analogical reasoning. (Tech. Rep. 1.) Minneapolis: Department of Psychology, University of Minnesota, 1972.

DeSoto, C. B. The predilection for single orderings. Journal of Abnormal and Social Psychology, 1961, 62, 16-23.

DeSoto, C. B., London, M., & Handel, S. Social reasoning and spatial paralogic. Journal of Personality and Social Psychology, 1965, 2, 513-521.

Donaldson, M. A study of children's thinking. London: Tavistock, 1963.

Dickstein, L. S. Effects of instructions and premise order on errors in syllogistic reasoning. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 376-384.

Dickstein, L. S. Differential difficulty of categorical syllogisms. Bulletin of the Psychonomic Society, 1976, 8, 330-332.

Dickstein, L. S. Error processes in syllogistic reasoning. Memory and Cognition, 1978, 6, 537-543.

Duncan, C. P. Recent research on human problem solving. Psychological Bulletin, 1959, 56, 397-429.

Duncker, K. A qualitative (experimental and theoretical) study of productive thinking (solving of comprehensible problems). Journal of Genetic Psychology, 1926, 33, 642-708.

Duncker, K. On problem-solving. Psychological Monographs, 1945, 58 (5), (Whole No. 270).

Egan, D. E. Accuracy and latency scores as measures of spatial information processing (Research report No. 1224). Pensacola, Florida: Naval Aerospace Medical Research Laboratories, 1976.

Egan, D. E., & Greeno, J. G. Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning, and problem solving. In L. W. Gregg (Ed.), Knowledge and cognition. Hillsdale, N.J.: Erlbaum, 1974.

Erickson, J. R. A set analysis theory of behavior in formal syllogistic reasoning tasks. In R. L. Solso (Ed.), Theories of cognitive psychology: The Loyola symposium. Hillsdale, N.J.: Erlbaum, 1974.

Erickson, J. R. Research on syllogistic reasoning. In R. Revlin & R. E. Mayer (Eds.), Human reasoning. New York: Wiley, 1978.

Ernst, G. W., & Newell, A. GPS: A case study in generality and problem-solving.

New York: Academic Press, 1969.

Esher, F. J. S., Raven, J. C., & Earl, C. J. C. Discussion on testing intellectual capacity in adults. Proceedings of the Royal Society of Medicine, 1942, 35, 779-785.

Evans, T. G. A program for the solution of geometric-analogy intelligence test questions. In M. Minsky (Ed.), Semantic information processing. Cambridge, Mass.: M.I.T. Press, 1968.

Ferguson, G. A. On learning and human ability. Canadian Journal of Psychology, 1954, 8, 95-112.

Feuerstein, R. The dynamic assessment of retarded performers: The learning potentia assessment device, theory, instruments, and techniques. Baltimore: University Park Press, 1979.

Feuerstein, R., Schlesinger, I. M., Shalom, H., & Narrol, H. The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques. Vol. II: LPAD analogies group test experiment. Studies in cognitive modifiability, Report No. 1. Jerusalem: Hadassah Wizo Canada Research Institute, 1972.

Fischhoff, B. Attribution theory and judgment under uncertainty. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), New directions in attribution research. Hillsdale, N.J.: Erlbaum, 1976.

Fleishman, E. A. The prediction of total task performance from prior practice on task components. Human Factors, 1965, 7, 18-27.

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. Journal of Experimental Psychology, 1955, 49, 301-312.

Frase, L. T. Belief, incongruity, and syllogistic reasoning. Psychological Reports, 1966, 18, 982. (a)

Frase, L. T. Validity judgments of syllogisms in relation to two sets of terms.

Journal of Educational Psychology, 1966, 57, 239-245. (b)

Frase, L. T. Associative factors in syllogistic reasoning. Journal of Experimental Psychology, 1968, 76, 407-412.

Gabriel, K. R. The simplex structure of the Progressive Matrices Test. British Journal of Statistical Psychology, 1954, 7, 9-14.

Gallagher, J. M., & Wright, R. J. Children's solution of verbal analogies: extension of Piaget's concept of reflexive abstraction. Paper presented at the annual meeting of the Society for Research in Child Development, New Orleans, 1977.

Gallagher, J. M., & Wright, R. J. Piaget and the study of analogy: structural analysis of items. In J. Magary (Ed.), Piaget and the helping professions (Vol. 8). Los Angeles: University of Southern California, 1979.

Garner, W. R. The processing of information and structure. Hillsdale, N.J.: Erlbaum, 1974.

Gentile, J. R., Kessler, D. K., & Gentile, P. K. Process of solving analogy items. Journal of Educational Psychology, 1969, 60, 494-502.

Gentile, J. R., Tedesco-Stratton, L., Davis, E., Lund, N. J., & Agunanne, B. A. Associative responding versus analogical reasoning by children. Intelligence, 1977, 1, 369-380.

Gick, M. L., & Holyoak, K. J. Analogical problem solving. Cognitive Psychology, in press.

Gilson, C., & Abelson, R. P. The subjective use of inductive evidence. Journal of Personality and Social Psychology, 1965, 2, 301-310.

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967.

Gollub, H. F., Rossman, B. B., & Abelson, R. P. Social inference as a function of the number of instances and consistency of information presented.

Journal of Personality and Social Psychology, 1973, 27, 19-33.

Greeno, J. G. Hobbits and orcs: Acquisition of a sequential concept. Cognitive Psychology, 1974, 6, 270-292.

Greeno, J. G. Natures of problem-solving abilities. In W. K. Estes (Ed.), Handbook of learning and cognitive processes (Vol. 5). Hillsdale, N.J.: Erlbaum, 1978.

Gregg, L. W. Internal representations of sequential concepts. In B. Kleinmuntz (Ed.), Concepts and the structure of memory. New York: Wiley, 1967.

Griggs, R. A. Logical processing of set inclusion relations in meaningful text. Memory and Cognition, 1976, 4, 730-740.

Griggs, R. A. Drawing inferences from set inclusion information given in text. In R. Revlin & R. E. Mayer (Eds.), Human reasoning. New York: V. H. Winston, 1978.

Grudin, J. Processes in verbal analogy solution. Journal of Experimental Psychology: Human Perception and Performance, 1980, 6, 67-74.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Guyote, M. J., & Sternberg, R. J. A transitive-chain theory of syllogistic reasoning. (NR 150-412 ONR Technical Report No. 5.) New Haven: Department of Psychology, Yale University, 1978.

Handel, S., DeSoto, C. B., & London, M. Reasoning and spatial representations. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 351-357.

Hayes, J. R., & Simon, H. A. Understanding written problem instructions. In L. W. Gregg (Ed.), Knowledge and cognition. Hillsdale, N.J.: Erlbaum, 1974.

Hayes, J. R., & Simon, H. A. Understanding complex task instructions. In D. Klahr (Ed.), Cognition and instruction. Hillsdale, N.J.: Erlbaum, 1976. (a)

Hayes, J. R., & Simon, H. A. The understanding process: Problem isomorphs.

Cognitive Psychology, 1976, 8, 165-190. (b)

Henle, M. On the relationship between logic and thinking. Psychological Review, 1962, 69, 366-378.

Henle, M., & Michael, M. The influence of attitudes on syllogistic reasoning. Journal of Social Psychology, 1956, 44, 115-127.

Hoffman, L. R., Burke, R. J., & Maier, N. R. F. Does training with differential reinforcement on similar problems help in solving a new problem? Psychological Reports, 1963, 13, 147-154.

Holyoak, K. J. Symbolic processes in mental comparisons. Unpublished doctoral thesis, Stanford University, 1976.

Holzman, T. G., Glaser, R., & Pellegrino, J. W. Process training derived from a computer simulation theory. Memory and Cognition, 1976, 4, 349-356.

Humphrey, G. Thinking: An introduction to its experimental psychology. New York: Wiley, 1951.

Hunt, E. Quote the raven? Nevermore! In L. W. Gregg (Ed.), Knowledge and cognition. Hillsdale, N.J.: Erlbaum, 1974.

Hunter, I. M. L. The solving of three term series problems. British Journal of Psychology, 1957, 48, 286-298.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-560.

Huttenlocher, J., Eisenberg, K., & Strauss, S. Comprehension: Relation between perceived actor and logical subject. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 334-341.

Huttenlocher, J., & Higgins, E. T. Adjectives, comparatives, and syllogisms. Psychological Review, 1971, 78, 487-504.

Huttenlocher, J., Higgins, E. T., Milligan, C., & Kauffman, B. The mystery of the "negative equative" construction. Journal of Verbal Learning and

Verbal Behavior, 1970, 9, 334-341.

Huttenlocher, J., & Strauss, S. Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 300-304.

Ingram, A. L., Pellegrino, J. W., & Glaser, R. Semantic processing in verbal analogies. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November, 1976.

Jacobs, P. I.,&Vandeventer, M. The learning and transfer of double-classification skills by first graders. Child Development, 1971, 42, 149-159. (a)

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills: A replication and extension. Journal of Experimental Child Psychology, 1971, 12, 140-157. (b)

Jacobs, P. I., & Vandeventer, M. Evaluating the teaching of intelligence. Educational and Psychological Measurement, 1972, 32, 235-248.

Janis, I., & Frick, F. The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. Journal of Experimental Psychology, 1943, 33, 73-77.

Jeffries, R., Polson, P. G., Razran, L., & Atwood, M. E. A process model for missionaries-cannibals and other river-crossing problems. Cognitive Psychology, 1977, 9, 412-440.

Johnson, D. M. The psychology of thought and judgment. New York: Harper & Row, 1955.

Johnson, D. M. Serial analysis of verbal analogy problems. Journal of Educational Psychology, 1962, 53, 86-88.

Johnson-Laird, P. N. The three-term series problem. Cognition, 1972, 1, 57-82.

Johnson-Laird, P. N. Models of deduction. In R. Falmagne (Ed.), Reasoning: Representation and process. Hilldale, N.J.: Erlbaum, 1975.

Johnson-Laird, P.N., & Steedman, M. The psychology of syllogisms. Cognitive Psychology, 1978, 10, 64-99.

Jones, M. R. From probability learning to sequential processing: A critical review. Psychological Bulletin, 1971, 76, 153-185.

Jöreskog, K. G. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 1969, 34, 183-202.

Jöreskog, K. G. A general method for analysis of covariance structure. Biometrika, 1970, 57, 239-251.

Jöreskog, K. G., & Sörbom, D. LISREL IV: Estimation of linear structural equation systems by maximum likelihood methods. Chicago: National Educational Resources, 1978.

Judson, A. J., Cofer, C. N., & Gelfand, S. Reasoning as an associative process: II. "Direction" in problem solving as a function of prior reinforcement of relevant responses. Psychological Reports, 1956, 2, 501-507.

Kaufman, H., & Goldstein, S. The effects of emotional value of conclusions upon distortions in syllogistic reasoning. Psychonomic Science, 1967, 7, 367-368.

Keating, D. P., & Caramazza, A. Effects of age and ability on syllogistic reasoning in early adolescence. Developmental Psychology, 1975, 11, 837-842.

Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium on motivation (Vol. 15). Lincoln: University of Nebraska Press, 1967.

Kelley, H. H. Attribution in social interaction. Morristown, N.J.: General Learning Press, 1972.

Klahr, D., & Wallace, J. G. The development of serial completion strategies: an information processing analysis. British Journal of Psychology, 1970, 61, 243-257.

Kling, R. E. A paradigm for reasoning by analogy. Artificial Intelligence, 1971, 2, 147-178.

Kodroff, J. K., & Roberge, J. J. Developmental analysis of the conditional reasoning abilities of primary-grade children. Developmental Psychology, 1975, 11, 21-28.

Köhler, W. The mentality of apes. London: Routledge & Kegan Paul, 1925.

Kosslyn, S. M., & Pomerantz, J. R. Imagery, propositions, and the form of internal representations. Cognitive Psychology, 1977, 9, 52-76.

Kotovsky, K., & Simon, H. A. Empirical tests of a theory of human acquisition of concepts for sequential events. Cognitive Psychology, 1973, 4, 399-424.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffreys (Ed.), Cerebral mechanisms in behavior: The Hixon Symposium. New York: Wiley, 1951.

Leeuwenberg, E. L. L. Quantitative specification of information in sequential patterns. Psychological Review, 1969, 76, 216-220.

Lefford, A. The influence of emotional subject matter on logical reasoning. Journal of General Psychology, 1946, 34, 127-151.

Levi, E. H. An introduction to legal reasoning. Chicago: University of Chicago Press, 1949.

Levinson, P. J., & Carpenter, R. L. An analysis of analogical reasoning in children. Child Development, 1974, 45, 857-861.

Linn, M. C. The role of intelligence in children's responses to instruction. Psychology in the Schools, 1973, 10, 67-75.

Lippman, M. Z. The influence of grammatical transform in a syllogistic reasoning task. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 424-430.

Luce, R. D. Individual choice behavior. New York: Wiley, 1959.

Luchins, A. S. Mechanization in problem solving. Psychological Monographs, 1942, 54, No. 6 (Whole No. 248).

Lunzer, E. A. Problems of formal reasoning in test situations. In P. H. Mussen (Ed.), European research in cognitive development. Monographs of the Society for Research in Child Development, 1965, 30(2, Serial No. 100), 19-46.

Lutkus, A. D., & Trabasso, T. Transitive inferences in preoperational retarded adolescents. American Journal of Mental Deficiency, 1974, 78, 599-606.

Lyon, D., & Slovic, P. Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychologica, 1976, 40, 287-298.

Maier, N. R. F. Reasoning in humans: I. On direction. Journal of Comparative Psychology, 1930, 10, 115-143.

Maier, N. R. F. Reasoning in humans: II. The solution of a problem and its appearance in consciousness. Journal of Comparative Psychology, 1931, 12, 181-194.

Maier, N. R. F. An aspect of human reasoning. British Journal of Psychology, 1933, 24, 144-155.

Maier, N. R. F. Reasoning in humans: III. The mechanisms of equivalent stimuli and of reasoning. Journal of Experimental Psychology, 1945, 35, 349-360.

Maier, N. R. F. What makes a problem difficult? In N. R. F. Maier (Ed.), Problem solving and creativity in individuals and groups. Belmont, California: Brooks/Cole, 1970.

Maier, N. R. F., & Burke, R. J. Test of the concept of "availability of functions" in problem solving. Psychological Reports, 1966, 19, 119-125.

Maier, N. R. F., & Janzen, J. C. Functional values as aids and distractors in problem solving. Psychological Reports, 1968, 22, 1021-1034.

Maltzman, I. Thinking: From a behavioristic point of view. Psychological Review, 1955, 62, 275-286.

Marcus, S., & Rips, L. J. Conditional reasoning. Journal of Verbal Learning and Verbal Behavior, 1979, 18, 199-224.

McGuire, W. J. A syllogistic analysis of cognitive relationships. In M. J. Rosenberg & C. I. Hovland (Eds.), Attitude organization and change. New Haven: Yale University Press, 1960.

Meer, B., Stein, M., & Geertsma, R. An analysis of the Miller Analogies Test for a scientific population. American Psychologist, 1955, 10, 33-34.

Mill, J. S. A system of logic, ratiocinative and inductive. London: J. W. Parker, 1843.

Miller, G. A. Images and models, similes and metaphors. In A. Ortony (Ed.), Metaphor and thought. New York: Cambridge University Press, 1979.

Miller Analogies Test manual. New York: The Psychological Corporation, 1970.

Morgan, J. J. B. Psychology. New York: Farrar & Rinehart, 1941.

Morgan, J. J., & Morton, J. T. The distortion of syllogistic reasoning produced by personal convictions. Journal of Social Psychology, 1944, 20, 39-59.

Mortensen, U. Models for some elementary problem solving processes. In A. Elithorn & D. Jones (Eds.), Artificial and human thinking. San Francisco: Jossey-Bass, 1973.

Moyer, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception and psychophysics, 1973, 13, 180-184.

Moyer, R. S., & Bayer, R. H. Mental comparison and the symbolic distance effect. Cognitive Psychology, 1976, 8, 228-246.

Mulholland, T. M., Pellegrino, J. W., & Glaser, R. Components of geometric analogy solution. Cognitive Psychology, in press.

Newell, A. You can't play 20 questions with nature and win. In W. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Nisbett, R. E., & Borgida, E. Attribution and the psychology of prediction. Journal of Personality and Social Psychology, 1975, 32, 932-943.

Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. Popular induction: Information is not necessarily informative. In J. S. Carroll & J. W. Payne (Eds.), Cognition and social behavior. Hillsdale, N.J.: Erlbaum, 1976.

Nisbett, R. E., & Ross, L. Human inferences: Strategies and shortcomings in social judgment. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Nisbett, R. E., & Wilson, T. D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Noble, C. E., Noble, J. L., & Alcock, W. T. Prediction of individual differences in human trial-and-error learning. Perceptual and Motor Skills, 1958, 8, 151-172.

Oppenheimer, J. R. Analogy in science. American Psychologist, 1956, 11, 127-135.

Osherson, D. N. Logical abilities in children (Vol. 2): Logical inference: Underlying operations. Hillsdale, N.J.: Erlbaum, 1974.

Osherson, D. N. Logical abilities in children (Vol. 3): Reasoning in adolescence: Deductive inference. Hillsdale, N.J.: Erlbaum, 1975.

Paris, S. G. Comprehension of language connectives and propositional logical relationships. Journal of Experimental Child Psychology, 1973, 16, 278-291.

Pellegrino, J. W., & Glaser, R. Components of inductive reasoning. In R. E. Snow, P.-A. Federico, & W. Montague (Eds.), Aptitude, learning, and instruction: Cognitive process analysis. Hillsdale, N.J.: Erlbaum, 1980.

Piaget, J. Une forme verbal de la comparison chez l'enfant. Archives de Psychologie, 1921, 141-172.

Piaget, J. Judgment and reasoning in the child. London: Routledge & Kegan Paul, 1928.

Piaget, J. The language and thought of the child. New York: Meridian Books Inc., 1955.

Piaget, J. Genetic epistemology. (Tr. E. Duckworth.) New York: Columbia University Press, 1970.

Piaget, J. (with Montangero, J., & Billeter, J.). Les correlats. L'Abstraction réfléchissante. Paris; Presses Universitaires de France, 1977.

Pintner, R. Contribution to "Intelligence and its measurement: A symposium." Journal of Educational Psychology, 1921, 12, 139-143.

Potts, G. R. Information processing strategies used in the encoding of linear orderings. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 727-740.

Potts, G. R. Storing and retrieving information about ordered relationships. Journal of Experimental Psychology, 1974, 103, 431-439.

Potts, G. R., & Scholz, K. W. The internal representation of a three-term series problem. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 439-452.

Psotka, J. Simplicity, symmetry, and syntely: Stimulus measures of binary pattern structure. Memory and Cognition, 1975, 3, 434-444.

Psotka, J. Syntely: Paradigm for an inductive psychology of memory, perception, and thinking. Memory and Cognition, 1977, 5, 553-560.

Pylyshyn, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 1973, 80, 1-24.

Quinton, G., & Fellows, B. 'Perceptual' strategies in the solving of three-

term series problems. British Journal of Psychology, 1975, 66, 69-78.

Raaheim, K. Problem solving and intelligence. Oslo: Universitetsforlaget, 1974.

Raven, J. C. Progressive Matrices: A perceptual test of intelligence, 1938, individual form. London: Lewis, 1938.

Raven, J. C. Guide to the standard progressive matrices. London: Lewis, 1960.

Ray, W. S. Complex tasks for use in human problem-solving research. Psychological Bulletin, 1955, 52, 134-149.

Reed, S. K. Pattern recognition and categorization. Cognitive Psychology, 1972, 3, 383-407.

Reed, S. K., & Abramson, A. Effect of the problem space on subgoal facilitation. Journal of Educational Psychology, 1976, 68, 243-246.

Reed, S. K., Ernst, G. W., & Banerji, R. The role of analogy in transfer between similar problem states. Cognitive Psychology, 1974, 6, 436-450.

Reitman, W. Cognition and thought. New York: Wiley, 1965.

Restle, F. Grammatical analysis of the prediction of binary events. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 17-25.

Restle, F. Theory of serial pattern learning: Structural trees. Psychological Review, 1970, 77, 481-495.

Restle, F. Serial patterns: The role of phrasing. Journal of Experimental Psychology, 1972, 92, 385-390.

Restle, F., & Brown, E. R. Serial pattern learning. Journal of Experimental Psychology, 1970, 83, 120-125. (a)

Restle, F., & Brown, E. R. Organization of serial pattern learning. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 4). New York: Academic Press, 1970. (b)

Revlin, R., & Leirer, V. O. The effects of personal biases on syllogistic reasoning:

Rational decision from personalized representations. In R. Revlin & R. E. Mayer (Eds.), Human reasoning. Washington: V. H. Winston, 1978.

Revlis, R. Syllogistic reasoning: Logical decisions from a complex data base. In R. Falmagne (Ed.), Reasoning: Representation and process. Hillsdale, N.J.: Erlbaum, 1975. (a)

Revlis, R. Two models of syllogistic reasoning: Feature selection and conversion. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 180-195. (b)

Richter, M. The theoretical interpretation of errors in syllogistic reasoning. Journal of Psychology, 1957, 43, 341-344.

Riley, C. A. The representation of comparative relations and the transitive inference task. Journal of Experimental Child Psychology, 1976, 22, 1-22.

Riley, C. A., & Trabasso, T. Comparatives, logical structures, and encoding in a transitive inference task. Journal of Experimental Child Psychology, 1974, 17, 187-203.

Rips, L. J., & Marcus, S. Suppositions and the analysis of conditional sentences. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1977.

Rips, L., Shoben, E., & Smith, E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Roberge, J. J., & Paulus, D. H. Developmental patterns for children's class and conditional reasoning abilities. Developmental Psychology, 1971, 4, 191-200.

Rosch, E. Human categorization. In N. Warren (Ed.), Advances in cross-cultural psychology (Vol. 1). London: Academic Press, 1977.

Rumelhart, D. E., & Abrahamson, A. A. A model for analogical reasoning. Cognitive Psychology, 1973, 5, 1-28.

Russell, D. H. Children's thinking. Boston: Ginn, 1956.

Saugstad, P. Problem-solving as dependent on availability of functions. British Journal of Psychology, 1955, 46, 191-198.

Saugstad, P. An analysis of Maier's pendulum problem. Journal of Experimental Psychology, 1957, 54, 168-179.

Saugstad, P. Availability of functions: A discussion of some theoretical aspects. Acta Psychologica, 1958, 14, 384-400.

Scriven, M. Maximizing the power of causal investigations: The modus operandi method. In G. V. Glass (Ed.), Evaluation studies: Review annual (Vol. 1). Beverly Hills: Sage Publications, 1976.

Sells, S. B. The atmosphere effect: An experimental study of reasoning. Archives of Psychology, 1936, 29, 3-72.

Shalom, H., & Schlesinger, I. M. Analogical thinking: A conceptual analysis of analogy tests. In R. Feuerstein, I. M. Schlesinger, H. Shalom, & H. Narrol (Eds.), Studies in cognitive modifiability (Report 1, Vol. II). Jerusalem: Hadassah Wizo Canada Research Institute, 1972.

Schank, R. C. Reminding and memory organization: An introduction to MOPs. (ARPA Technical Report #170.) Department of Computer Science, Yale University, 1979.

Shaver, P., Pierson, L., & Lang, S. Converging evidence for the functional significance of imagery in problem solving. Cognition, 1974, 3, 359-375.

Shepard, R. N., & Arabie, P. Additive clustering: Representation of similarities as combinations of discrete overlapping properties. Psychological Review, 1979, 86, 87-123.

Simon, H. A. Complexity and the representation of patterned sequences of symbols. Psychological Review, 1972, 79, 369-382.

Simon, H. A. The functional equivalence of problem solving skills. _Cognitive Psychology_, 1975, _7_, 268-288.

Simon, H. A. Identifying basic abilities underlying intelligent performance of complex tasks. In L. Resnick (Ed.), _The nature of intelligence_. Hillsdale, N.J.: Erlbaum, 1976.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. _Psychological Review_, 1963, _70_, 534-546.

Simon, H. A., & Lea, G. Problem solving and rule induction: A unified view. In L. W. Gregg (Ed.), _Knowledge and cognition_. Hillsdale, N.J.: Erlbaum, 1974.

Simon, H. A., & Newell, A. Thinking processes. In D. H. Krant, R. D. Luce, Atkinson, R.C. & Suppes, P. (Eds.), _Contemporary developments in mathematical psychology_ (Vol. 1). San Francisco: Freeman, 1974.

Simon, H. A., & Reed, S. K. Modeling strategy shifts in a problem-solving task. _Cognitive Psychology_, 1976, _8_, 86-97.

Simon, H. A., & Sumner, R. K. Pattern in music. In B. Kleinmuntz (Ed.), _Formal representation of human judgment_. New York: Wiley, 1968.

Simpson, M. E., & Johnson, D. M. Atmosphere and conversion errors in syllogistic reasoning. _Journal of Experimental Psychology_, 1966, _72_, 197-200.

Skyrms, B. _Choice and chance_ (2nd ed.). Encino, California: Dickenson, 1975.

Smedslund, J. The concept of correlation in adults. _Scandinavian Journal of Psychology_, 1963, _4_, 165-173.

Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A feature model for semantic decisions. _Psychological Review_, 1974, _81_, 214-241.

Spearman, C. 'General intelligence,' objectively determined and measured. _American Journal of Psychology_, 1904, _15_, 201-293.

Spearman, C. The nature of 'intelligence' and the principles of cognition. London: Macmillan, 1923.

Spearman, C. The abilities of man. New York: Macmillan, 1927.

Staudenmayer, H. Understanding conditional reasoning with meaningful propositions. In R. Falmagne (Ed.), Reasoning: Representation and process. Hillsdale, N.J.: Erlbaum, 1975.

Staudenmayer, H., & Bourne, L. E., Jr. Learning to interpret conditional sentences: A developmental study. Developmental Psychology, 1977, 13, 616-623.

Sternberg, R. J. Component processes in analogical reasoning. Psychological Review, 1977, 84, 353-378. (a)

Sternberg, R. J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, N.J.: Erlbaum, 1977. (b)

Sternberg, R. J. Componential investigations of human intelligence. In A. Lesgold, J. Pellegrino, S. Fokkema, & R. Glaser (Eds.), Cognitive psychology and instruction. New York: Plenum, 1978. (a)

Sternberg, R. J. Isolating the components of intelligence. Intelligence, 1978, 2, 117-128. (b)

Sternberg, R. J. Intelligence research at the interface between differential and cognitive psychology. Intelligence, 1978, 2, 195-222. (c)

Sternberg, R. J. Developmental patterns in the encoding and combination of logical connectives. Journal of Experimental Child Psychology, 1979, 28, 469-498. (a)

Sternberg, R. J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230. (b)

Sternberg, R. J. Six authors in search of a character: A play about intelligence tests in the year 2000. Intelligence, 1979, 3, 281-291. (c)

Sternberg, R. J. Stalking the I.Q. quark. Psychology Today, 1979, 13, 42-54. (d)

Sternberg, R. J. Componentman as vice-president: A reply to Pellegrino and Lyon's analysis of "The components of a componential analysis." Intelligence, 1980, 4, 83-95. (a)

Sternberg, R. J. The development of linear syllogistic reasoning. Journal of Experimental Child Psychology, 1980, 29, 340-356. (b)

Sternberg, R. J. A proposed resolution of curious conflicts in the literature on linear syllogistic reasoning. In R. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum, 1980. (c)

Sternberg, R. J. Representation and process in linear syllogistic reasoning. Journal of Experimental Psychology: General, 1980, 109, 119-159. (d)

Sternberg, R. J. Sketch of a componential subtheory of human intelligence. Behavioral and Brain Sciences, in press.

Sternberg, R. J., & Davis, J. C. Student perceptions of Yale and its competitors. College and University, 1978, 53, 262-278.

Sternberg, R. J., & Gardner, M. K. Unities in inductive reasoning. (NR 150-412 ONR Technical Report #18.) New Haven: Department of Psychology, Yale University, 1979.

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. Peoples' conceptions of intelligence. Manuscript submitted for publication, 1980.

Sternberg, R. J., Ketron, J. L., & Powell, J. S. Componential approaches to the training of intelligence. Intelligence, in press.

Sternberg, R. J., & Nigro, G. Developmental patterns in the solution of verbal analogies. Child Development, 1980, 51, 27-38.

Sternberg, R. J., & Rifkin, B. The development of analogical reasoning processes. Journal of Experimental Child Psychology, 1979, 27, 195-232.

Sternberg, R. J., Tourangeau, R., & Nigro, G. Metaphor, induction, and social

policy: The convergence of macroscopic and microscopic views. In A. Ortony (Ed.), Metaphor and thought. New York: Cambridge University Press, 1979.

Sternberg, R. J., & Tulving, E. The measurement of subjective organization in free recall. Psychological Bulletin, 1977, 84, 539-556.

Sternberg, R. J., & Turner, M. E. Components of syllogistic reasoning. Acta Psychologica, in press.

Sternberg, R. J., & Weil, E. M. An aptitude-strategy interaction in linear syllogistic reasoning. Journal of Educational Psychology, 1980, 72, 226-234.

Stevens, S. S. Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951.

Taplin, J. E. Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 219-225.

Taplin, J. E., & Staudenmayer, H. Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 530-542.

Taplin, J. E., Staudenmayer, H., & Taddonio, J. L. Developmental changes in conditional reasoning: Linguistic or logical? Journal of Experimental Child Psychology, 1974, 17, 360-373.

Taylor, S. E., & Fiske, S. T. Salience, attention, and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 11). New York: Academic Press, 1978.

Terman, L. M. Contribution to "Intelligence and its measurement: A symposium." Journal of Educational Psychology, 1921, 12, 127-133.

Terman, L. M., & Merrill, M. A. Measuring intelligence. Boston: Houghton-Mifflin, 1937.

Thomas, J. C., Jr. An analysis of behavior in the hobbits-orcs problem. Cognitive Psychology, 1974, 6, 257-269.

Thurstone, L. L. Primary mental abilities. Chicago: University of Chicago Press, 1938.

Thurstone, L. L., & Thurstone, T. G. SRA Primary Mental Abilities. Chicago: Science Research Associates, 1962.

Tinsley, H. E. A., & Davis, R. J. The equivalence of semantic and figural test presentation of the same items. The Center for the Study of Organizational Performance and Human Effectiveness. Technical Report No. 3004. Minneapolis: University of Minnesota, 1972.

Tourangeau, R., & Sternberg, R. J. Aptness in metaphor. Cognitive Psychology, in press.

Trabasso, T.. Representation, memory, and reasoning: How do we make transitive inferences? In A. D. Pick (Ed.), Minnesota symposia on child psychology (Vol. 9). Minneapolis: University of Minnesota Press, 1975.

Trabasso, T., & Riley, C. A. On the construction and use of representations involving linear order. In R. L. Solso (Ed.), Information processing and cognition: The Loyola Symposium. Hillsdale, N.J.: Erlbaum, 1975.

Trabasso, T., Riley, C. A., & Wilson, E. G. The representation of linear order and spatial strategies in reasoning: A developmental study. In R. Falmagne (Ed.), Reasoning: Representation and process. Hillsdale, N.J.: Erlbaum, 197

Tulving, E. Subjective organization and effects of repetition in multi-trial free-recall learning. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 193-197.

Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

Tversky, A., & Kahneman, D. Causal schemata in judgments under uncertainty.

In M. Fishbein (Ed.), Progress in social psychology. Hilldale, N.J.: Erlbaum, 1977.

Underwood, B. J. An orientation for research on thinking. Psychological Review, 1952, 59, 209-220.

van de Geer, J. P. A psychological study of problem solving. Haarlem: Uitgeverij De Toorts, 1957.

Vinacke, W. E. The psychology of thinking. New York: McGraw-Hill, 1952.

Vitz, P. C., & Todd, T. C. A model of learning for simple repeating binary patterns. Journal of Experimental Psychology, 1967, 75, 108-117.

Vygotsky, L. S. Thought and language. (Edited and translated by E. Hanfmann & G. Vakar.) Cambridge, Mass.: M.I.T. Press, 1962.

Wason, P. C. 'On the failure to eliminate hypotheses...'--A second look. In P. C. Wason & P. N. Johnson-Laird (Eds.), Thinking and reasoning. Harmondsworth, England: Penguin, 1968.

Wason, P. C., & Johnson-Laird, P. N. Psychology of reasoning: Structure and content. London: B. T. Batsford, 1972.

Whitely, S. E. Types of relationships in reasoning by analogy. Unpublished doctoral dissertation, University of Minnesota, 1973.

Whitely, S. E. Solving verbal analogies: Some cognitive components of intelligence test items. Journal of Educational Psychology, 1976, 68, 234-242.

Whitely, S. E. Information-processing on intelligence test items. Some response components. Applied Psychological Measurement, 1977, 1, 465-476.

Whitely, S. E. Latent trait models in the study of intelligence. Intelligence, 1979, 4, 97-132. (a)

Whitely, S. E. Modeling aptitude test validity from cognitive components. (NIE 79-2 Technical Report.) Lawrence, Kansas: Department of Psychology, University of Kansas, 1979. (b)

Whitely, S. E. Multicomponent latent trait models for information-processes

on ability tests. (NIE 79-1 Technical Report.) Lawrence Kansas: Department of Psychology, University of Kansas, 1979. (c)

Whitely, S. E., & Barnes, G. M. The implications of processing event sequences for theories of analogical reasoning. Memory and Cognition, 1979, 1, 323-331.

Whitely, S. E., & Dawis, R. A cognitive intervention for improving the estimate of latent ability measured from analogy items. The Center for the Study of Organizational Performance and Human Effectiveness. Technical Report No. 3010. Minneapolis: University of Minnesota, 1973.

Whitely, S. E., & Dawis, R. V. The effects of cognitive intervention on latent ability measured from analogy items. Journal of Educational Psychology, 1974, 66, 710-717.

Wilkins, M. C. The effect of changed material on ability to do formal syllogistic reasoning. Archives of Psychology, 1923, 16, No. 102.

Williams, D. S. Computer program organization induced from problem examples. In H. A. Simon & L. Siklossy (Eds.), Representation and meaning: Experiments with information processing systems. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Willner, A. An experimental analysis of analogical reasoning. Psychological Reports, 1964, 15, 479-494.

Wilson, W. R. The effect of competition on the speed and accuracy of syllogistic reasoning. Journal of Social Psychology, 1965, 65, 27-32.

Winston, P. H. Learning structural descriptions from examples. Artificial Intelligence Laboratory. AI TR-231. Cambridge, Mass.: M.I.T., 1970.

Wood, D. J. The nature and development of problem-solving strategies. Unpublished doctoral thesis, University of Nottingham, 1969.

Wood, D., Shotter, J., & Godden, D. An investigation of the relationships between problem solving strategies, representation and memory. Quarterly Journal of Experimental Psychology, 1974, 26, 252-257.

Woodworth, R. S., & Schlosberg, H. Experimental psychology. New York: Holt, Rinehart, & Winston, 1954.

Woodworth, R. S., & Sells, S. B. An atmosphere effect in formal syllogistic reasoning. Journal of Experimental Psychology, 1935, 18, 451-460.

Footnote

Preparation of this chapter was supported by Contract N0001478C0025 from the Office of Naval Research to Robert J. Sternberg.

Technical Reports Presently in this Series

NR 150-412, ONR Contract N0001478C0025

No.	Name	Published Reference
1	Intelligence Research at the Interface between Differential and Cognitive Psychology. January, 1978.	Sternberg, R. J. Intelligence research at the interface between differential and cognitive psychology. Intelligence 1978, 2, 195-222.
2	Isolating the Components of Intelligence. January, 1978.	Sternberg, R. J. Isolating the component of intelligence. Intelligence, 1978, 2, 117-128.
3	Deductive Reasoning. January, 1978.	Sternberg, R. J., Guyote, M. J., & Turner, M. E. Deductive reasoning. In R. E. Snow, P.-A. Federico, & W. Montague (Eds.), Aptitude, learning and instruction: Cognitive process analysis (Vol. 1). Hillsdale, N.J.: Erlbaum, 1980.
4	Toward a Unified Componential Theory of Human Reasoning. April, 1978.	Sternberg, R. J. Toward a unified componential theory of human intelligence: I. Fluid ability. In M. Friedman, J. Das, & N. O'Connor (Eds.), Intelligence and learning. New York: Plenum, 1980.
5	A Transitive-Chain Theory of Syllogistic Reasoning. April, 1978.	UNPUBLISHED TO DATE
6	Components of Syllogistic Reasoning. April, 1978.	Sternberg, R. J., & Turner, M. E. Components of syllogistic reasoning. Acta Psychologica, in press.
7	Metaphor, Induction, and Social Policy: The Convergence of Macroscopic and Microscopic Views. April, 1978.	Sternberg, R. J., Tourangeau, R., & Nigro, G. Metaphor, induction, and social policy: The convergence of macroscopic and microscopic views. In A. Ortony (Ed.), Metaphor and thought New York: Cambridge University Press, 1979.
8	A Proposed Resolution of Curious Conflicts in the Literature on Linear Syllogisms. June, 1978.	Sternberg, R. J. A proposed resolution of curious conflicts in the literature on linear syllogisms. In R. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum, 1980.
9	The Nature of Mental Abilities. June, 1978.	Sternberg, R. J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230.

Technical Reports Presently in this Series

NR 150-412

Page 2

No.	Name	Published Reference
10	Psychometrics, Mathematical Psychology, and Cognition: Confessions of a Closet Psychometrician. June, 1978.	UNPUBLISHABLE.
11	Understanding and Appreciating Metaphors. June, 1978.	UNPUBLISHED TO DATE.
12.	Representation and Process in Transitive Inference. October, 1978.	Sternberg, R. J. Representation and process in linear syllogistic reasoning. Journal of Experimenta Psychology: General, 1980, 109, 119-159.
13	Aptness in Metaphor. October, 1978.	Tourangeau, R., & Sternberg, R. J. Aptness in metaphor. Cognitive Psychology, in press.
14	Contrasting Conceptions of Intelligence and their Educational Implications. November, 1978.	Sternberg, R. J. Factor theories of intelligence are all right almost. Educational Researcher, in press.
15	An Aptitude-Strategy Interaction in Linear Syllogistic Reasoning. April, 1979.	Sternberg, R. J., & Weil, E. M. An aptitude-strategy interaction i linear syllogistic reasoning. Journal of Educational Psychology, 1980, 72, 226-234.
16	Intelligence Tests in the Year 2000: What Forms will they Take and what Purposes will they Serve? April, 1979.	Sternberg, R. J. Six authors in search of a character: A play about intelligence tests in the year 2000. Intelligence, 1979, 3, 281-291.
17	New Views on IQs: A Silent Revolution of the 70s. April, 1979.	Sternberg, R. J. Stalking the I.Q. quark. Psychology Today, 1979, 13 42-54.
18	Unities in Inductive Reasoning. October, 1979.	UNPUBLISHED TO DATE.
19	Components of Human Intelligence. October, 1979.	Sternberg, R. J. Sketch of a componential subtheory of human intelligence. Behavioral and Brain Sciences, in press.
20	The Construct Validity of Aptitude Tests: An Information-Processing Assessment. October, 1979.	Sternberg, R. J. The construct vali ty of aptitude tests: An information-processing assessment. In

Technical Reports Presently in this Series

NR 150-412

No.	Name	Published Reference
20 (Continued)		A. P. Maslow, R. H. McKillup, & M. Thatcher (Eds.), Construct validity in psychological measurement. Princeton: Educational Testing Service, in press.
21	Evaluation of Evidence in Causal Inference. October, 1979.	Schustack, M. W., & Sternberg, R. J. Evaluation of evidence in causal inference. Journal of Experimental Psychology: General, in press.
22	Componential Approaches to the Training of Intelligent Performance. April, 1980.	Sternberg, R. J., Ketron, J. L., & Powell, J. S. Componential approaches to the training of intelligent performance. Intelligence, in press.
23	Intelligence and Nonentrenchment. April, 1980.	UNPUBLISHED TO DATE.
24	Reasoning, Problem Solving, and Intelligence. April, 1980.	Sternberg, R. J. Reasoning, problem solving, and intelligence. In R. J. Sternberg (Ed) Handbook of human intelligence. New York: Cambridge University Press, in press.
25	Claims, Counterclaims, and Components: A Countercritique of Componential Analysis. June, 1980.	Sternberg, R. J. Claims, counterclaims, and components: A countercritique of componential analysis. Behavioral and Brain Sciences, in press.

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Meryl S. Baker
NPRDC
Code P309
San Diego, CA 92152

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Chief of Naval Education and Training
Liason Office
Air Force Human Resource Laboratory
Flying Training Division
WILLIAMS AFB, AZ 85224

1 Dr. Larry Dean, LT, MSC, USN
Psychology Department
Naval Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 LT Steven D. Harris, MSC, USN
Code 6021
Naval Air Development Center
Warminster, Pennsylvania 18974

1 Dr. Patrick R. Harrison
Psychology Course Director
LEADERSHIP & LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 CDR Charles W. Hutchins
Naval Air Systems Command Hq
AIR-340F
Navy Department
Washington, DC 20361

1 CDR Robert S. Kennedy
Head, Human Performance Sciences
Naval Aerospace Medical Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 Dr. George Moeller
Head, Human Factors Dept.
Naval Submarine Medical Research Lab
Groton, CN 06340

1 Dr William Montague
Navy Personnel R&D Center
San Diego, CA 92152

1 Library
Naval Health Research Center
P. O. Box 85122
San Diego, CA 92138

1 Naval Medical R&D Command
Code 44
National Naval Medical Center
Bethesda, MD 20014

1 CAPT Paul Nelson, USN
Chief, Medical Service Corps
Bureau of Medicine & Surgery (MED-23)
U. S. Department of the Navy
Washington, DC 20372

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

Navy

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

1 Office of Naval Research
Code 441
800 N. Quincy Street
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research, Development, and Studies Branc
(OP-102)
Washington, DC 20350

1 Captain Donald F. Parker, USN
Commanding Officer
Navy Personnel R&D Center
San Diego, CA 92152

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 DR. RICHARD A. POLLAK
ACADEMIC COMPUTING CENTER
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Roger W. Remington, Ph.D
Code L52
NAMRL
Pensacola, FL 32508

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

Navy

1 Mr. Arnold Rubenstein
Naval Personnel Support Technology
Naval Material Command (08T244)
Room 1044, Crystal Plaza #5
2221 Jefferson Davis Highway
Arlington, VA 20360

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Sam Schiflett, SY 721
Systems Engineering Test Directorate
U.S. Naval Air Test Center
Patuxent River, MD 20670

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group (TAEG)
Dept. of the Navy
Orlando, FL 32813

1 W. Gary Thomson
Naval Ocean Systems Center
Code 7132
San Diego, CA 92152

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Navy

1 Mr John H. Wolfe
Code P310
U. S. Navy Personnel Research and Development Center
San Diego, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQ USAREUE & 7th Army
ODCSOPS
USAAREUE Director of GED
APO New York 09403

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. Ronald G. Hughes
AFHRL/OTR
Williams AFB, AZ 85224

1 Dr. Ross L. Morgan (AFHRL/LR)
Wright -Patterson AFB
Ohio 45433

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

2 3700 TCHTW/TTGH Stop 32
Sheppard AFB, TX 76311

1 Jack A. Thorpe, Maj., USAF
Naval War College
Providence, RI 02846

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 Chief, Psychological Reserch Branch
U. S. Coast Guard (G-P-1/2/TP42)
Washington, DC 20593

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Joseph I. Lipson
SEDR W-638
National Science Foundation
Washington, DC 20550

1 William J. McLaurin
Rm. 301, Internal Revenue Service
2221 Jefferson Davis Highway
Arlington, VA 22202

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. John R. Anderson
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Anderson, Thomas H., Ph.D.
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champiagn, IL 61820

1 Dr. John Annett
Department of Psychology
University of Warwick
Coventry CV4 7AL
ENGLAND

1 DR. MICHAEL ATWOOD
SCIENCE APPLICATIONS INSTITUTE
40 DENVER TECH. CENTER WEST
7935 E. PRENTICE AVENUE
ENGLEWOOD, CO 80110

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

1 Dr. Patricia Baggett
Department of Psychology
University of Denver
University Park
Denver, CO 80208

1 Mr Avron Barr
Department of Computer Science
Stanford University
Stanford, CA 94305

Non Govt

1 Dr. Jackson Beatty
Department of Psychology
University of California
Los Angeles, CA 90024

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Lyle Bourne
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1 Dr. John S. Brown
XEROX Palo Alto Research Center
3333 Coyote Road
Palo Alto, CA 94304

1 Dr. Bruce Buchanan
Department of Computer Science
Stanford University
Stanford, CA 94305

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. Pat Carpenter
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Non Govt

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. William Chase
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Micheline Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

1 Dr. William Clancey
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, Ma 02138

Non Govt

1 Dr. Lynn A. Cooper
Department of psychology
Uris Hall
Cornell University
Ithaca, NY 14850

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

1 Dr. Kenneth B. Cross
Anacapa Sciences, Inc.
P.O. Drawer Q
Santa Barbara, CA 93102

1 Dr. Emmanuel Donchin
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. Hubert Dreyfus
Department of Philosophy
University of California
Berkely, CA 94720

1 LCOL J. C. Eggenberger
DIRECTORATE OF PERSONNEL APPLIED RESEARC
NATIONAL DEFENCE HQ
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A OK2

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Ed Feigenbaum
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

Non Govt

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Alinda Friedman
Department of Psychology
University of Alberta
Edmonton, Alberta
CANADA T6G 2E9

1 Dr. R. Edward Geiselman
Department of Psychology
University of California
Los Angeles, CA 90024

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Daniel Gopher
Industrial & Management Engineering
Technion-Israel Institute of Technology
Haifa
ISRAEL

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

Non Govt

1 Dr. Harold Hawkins
Department of Psychology
University of Cregon
Eugene OR 97403

1 Dr. Barbara Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. Frederick Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. James R. Hoffman
Department of Psychology
University of Delaware
Newark, DE 19711

1 Glenda Greenwald, Ed.
"Human Intelligence Newsletter"
P. O. Box 1163
Birmingham, MI 48012

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

3 Journal Supplement Abstract Service
American Psychological Association
1200 17th Street N.W.
Washington, DC 20036

1 Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

Non Govt

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. David Kieras
Department of Psychology
University of Arizona
Tuscon, AZ 85721

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Jill Larkin
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Mark Miller
Computer Science Laboratory
Texas Instruments, Inc.
Mail Station 371, P.O. Box 225936
Dallas, TX 75265

Non Govt

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

1 Dr. Donald A Norman
Dept. of Psychology C-009
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. Seymour A. Papert
Massachusetts Institute of Technology
Artificial Intelligence Lab
545 Technology Square
Cambridge, MA 02139

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 Dr. Martha Polson
Department of Psychology
University of Colorado
Boulder, CO 80302

1 DR. PETER POLSON
DEPT. OF PSYCHOLOGY
UNIVERSITY OF COLORADO
BOULDER, CO 80309

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

Non Govt

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Fred Reif
SESAME
c/o Physics Department
University of California
Berkely, CA 94720

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 Dr. Irwin Sarason
Department of Psychology
University of Washington
Seattle, WA 98195

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 Dr. Alan Schoenfeld
Department of Mathematics
Hamilton College
Clinton, NY 13323

Non Govt

1 Committee on Cognitive Research
% Dr. Lonnie R. Sherrod
Social Science Research Council
605 Third Avenue
New York, NY 10016

1 Robert S. Siegler
Associate Professor
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Kathryn T. Spoehr
Department of Psychology
Brown University
Providence, RI 02912

1 DR. ALBERT STEVENS
BOLT BERANEK & NEWMAN, INC.
50 MOULTON STREET
CAMBRIDGE, MA 02138

1 Dr. David Stone
ED 236
SUNY, Albany
Albany, NY 12222

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

Non Govt

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. John Thomas
IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

1 DR. PERRY THORNDYKE
THE RAND CORPORATION
1700 MAIN STREET
SANTA MONICA, CA 90405

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

1 Dr. William R. Uttal
University of Michigan
Institute for Social Research
Ann Arbor, MI 48106

Non Govt

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 Dr. Keith T. Wescourt
Information Sciences Dept.
The Rand Corporation
1700 Main St.
Santa Monica, CA 90406

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Dr. Christopher Wickens
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. J. Arthur Woodward
Department of Psychology
University of California
Los Angeles, CA 90024

END